University of Sussex

Simulating the 21-cm Signal during the Cosmic Dawn

Hannah Ross

Submitted for the degree of Doctor of Philosophy

University of Sussex

July 2018

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Hannah Ross

UNIVERSITY OF SUSSEX

Hannah Ross, Doctor of Philosophy

Simulating the 21-cm Signal from the Cosmic Dawn

Abstract

The anticipated radio telescope SKA is expected to detect the 21-cm signal from the Cosmic Dawn, allowing us to probe the astrophysical processes of this previously unobserved era. The 21-cm differential brightness temperature fluctuations from the Cosmic Dawn are driven by early inhomogeneous heating of the neutral intergalactic medium and variations in Lyman-alpha photon density. Inhomogeneous heating is driven by high energy, X-ray photons which have long mean free paths and thus penetrate deep into the neutral intergalactic medium. Lyman-α fluctuations depend on the soft, UV photons from these sources redshifting into Lyman-α resonance. In this thesis I present a large-volume (349 Mpc comoving) suite of fully numerical radiative transfer simulations of this epoch. The simulations include the effects of helium ionisation, secondary ionisations and multi-frequency heating in order to include different types of X-ray sources (high mass X-ray binaries sources and QSO sources) in addition to black body stellar sources and Lyman-alpha fluctuations, which are added as a post-processing step.

In our simulations X-ray sources are able to contribute significantly to early heating of the neutral IGM. Different X-ray models produce varying lengths and morphologies of the transition from absorption to emission. When the results are smoothed to the expected resolution of SKA1-Low the mean, rms, skewness, kurtosis and power spectra of the 21-cm differential brightness temperature are notably different for each X-ray model. These rms fluctuations for each heating model are well above the expected noise for deep integrations which suggests direct imaging of X-ray heating during the Cosmic Dawn should be possible. The presence of QSOs greatly affects the non-Gaussianity, suggesting higher order statistics may be a good observational probe of rare X-ray sources. This effect is decreased if the Lyman-α background is built up late. We conclude by discussing ongoing and future work on the topic.

Acknowledgements

Firstly I would like to thank my supervisor, Ilian Iliev, for his guidance throughout my PhD. I have greatly appreciated not only your help in developing my skills as a researcher but also your advice and support with regards to my future career. I would also like to thank Garrelt Mellema for being my unofficial second supervisor! Working with you and your group in Stockholm has been a wonderful experience that certainly played a part in my decision to continue in academia. In fact, I feel that I have had not two but three supervisors due to the huge amount of effort and time put aside by Keri Dixon to help me with my PhD research. You not only offered vital technical help with my work but also helped teach me how to approach and solve daunting problems. To all the people mentioned in this paragraph, I hope that this is only the beginning of our collaboration.

In my time at Sussex I have met many amazing, kind and highly entertaining individuals. I would especially like to thank David Sullivan. The time you put into teaching me about cosmological simulations at the beginning of my PhD and your ongoing help with computational woes was invaluable. Also, thank you for not letting me eat too many healthy office snacks and for 'fat Thursdays'. My other officemates Alex, Chaichalit (Boyd), Azizah, Antonio, Sunayana, Carlos and Michaelle also deserve a mention for providing a not too distracting but entertaining working environment during my PhD. Benoit, thank you for all the tea breaks that stopped me working too hard and Dániel, thank you for all the late dinners when I was working too hard! Thank you to Steven for making sure everyone made it to the pub on Fridays (and sorry that I was rarely on time) and also thanks to Luke for introducing me to Scuba (and possibly causing my eye to need surgery by doing so). I would also like to thank (in no particular order and apologies to whoever I inevitably forget) Scott, Mateja, Ciaran, Pippa, Sean, Rose, Ridwan, Michaela, Eliza, Chris, and outside of the department Jonathan, Caroline and Raph for making my time in Brighton so fun!

From Stockholm University I would like to thank Raghunath Ghara for all the help and guidance during my time in Sweden, I'm sure that your patience and efficiency will make you a wonderful supervisor one day. I would also like to thank Sambit Giri for being a fantastic collaborator and even better friend. Your enthusiasm about science regularly reminds me why I am pursuing a career in research. My time in Stockholm wouldn't have been the same without such a warm and welcoming department, I would like to thank the following people: Himanshu, Armin, Kai-Yan, Emir, Aravind, Alex, Ruben, Veronica, Anders, Per, Illa, Suhail, and Matteo for all the fikas, chai breaks and evenings in the pub. Last but by no means least, I would like to thank Cecilia Syltern and all of my 'Swedish family' for welcoming me to Sweden and making me feel it was my second home!

Of course, I wouldn't be where I am today without the help and support of my family. Thank you to my parents, Joanne and David Ross, for always allowing my curiosity of the world around me to grow. Everything from the trips science museum as a small child to taking me to astrophysics talks when I was a teenager to simply always answering my questions, no matter how abstract, paved the way for me to reach where I am today. Thank you siblings: Freya, Wilfred and Oliver, for supporting me when things were bad, distracting me from my work when I needed it and helping me to see the funny side when things went wrong. I would also like to thank my grandparents: Dorothy and John Young, Anne and John Bore and Morgan Ross. Thank you to Ángelines Alonso Moraga for being an excellent role model and to Dulce Muñoz Alonso for always being so sweet!

Finally, I would like to thank my long term partner Alberto Muñoz Alonso. You have been with me every step of the way of this PhD, from application to submission, and have made these last four and a half years an incredible part of my life. You are my inspiration, my most valued friend and the love of my life. I dedicate this thesis to you.

"The truth, as always, will be far stranger."

Arthur C. Clarke (A space odyssey, 2001)

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Conceivably the most captivating question to astrophysicists is how the immensely complex structures we observe around us today evolved from the largely homogeneous Universe present after the Big Bang. While the overarching mechanism of structure formation is reasonably well understood, there are still vast gaps in our knowledge of how our Universe evolved. Perhaps the most significant of these is our lack of insight into the last global phase transition our Universe underwent: the Epoch of Reionization (EoR). The EoR refers to the cosmological epoch in which the early luminous sources almost completely ionized the initially neutral intergalactic medium (IGM).

While the EoR has been constrained by observations, it is doubtful as to whether or not it has been directly detected (more on this in section 1.4.4). The most promising observational probe of this time is the redshifted 21-cm signal between the hyperfine transition from the triplet to singlet ground state in hydrogen atoms, which makes up almost all baryonic matter from this time onwards. This signal is predicted to incorporate vast amounts of information about the astrophysical conditions of the time. For a comprehensive review on this topic see Furlanetto et al. (2006).

Currently, radio interferometers LOFAR[1], PAPER[2] GMRT[3] and MWA[4] are attempting to make a detection of this signal. These interferometers have not yet made a direct detection, however, they have placed constraints on the fluctuations of the signal (see section 1.4.4 for more details). Upcoming radio interferometry experiments HERA[5] (DeBoer et al., 2017)

[1] http://www.lofar.org/
[2] http://eor.berkeley.edu/
[3] http://www.gmrt.ncra.tifr.res.in/
[4] http://mwatelescope.org/
[5] http://reionization.org/

and SKA[6] (Mellema et al., 2013) are expected to fill this large gap in our observations and enable us to understand the astrophysics of this time.

The Cosmic Dawn (CD), occasionally known as the Epoch of Heating, encompasses the first stages of the EoR where significant ionization had not yet begun but long-range X-ray heating from the first sources was underway. The expected 21-cm signal from the later stages EoR has been considered in great detail due to the fact that lower redshifts are easier to detect but less attention has been paid to this earlier phase. The 21-cm signal from this time offers a unique glimpse into the properties of the first sources and their remnants via the heating they enact on the surrounding neutral IGM, information complementary to that from the 21-cm signal from later times of the EoR.

The contents of this Chapter provide a brief review of the cosmology and astrophysics related to the EoR. Section 1.1 outlines our understanding of the high redshift Universe and the cosmology behind its evolution. Section 1.2 focusses on the possible natures of the first sources and the effects they may have. In section 1.3 the physics behind and the evolution of the 21-cm signal that will be observed by SKA is outlined in detail. Observational constraints are discussed along with upcoming experiments, in particular, SKA in section 1.4. Finally we conclude the introduction in section 1.5 with an introduction to fully numerical simulations.

1.1 A Brief History of the Universe

The Big Bang model is the widely-accepted accepted model for the birth and growth of our Universe. Cold dark matter with a cosmological constant (ΛCDM) has been immensely successful in describing how the Universe evolved while matching observations. ΛCDM postulates that the Universe began as a singularity which expanded so rapidly that quantum fluctuations were stretched sufficiently to seed the gravitational collapse of the structure we see on all scales today. Here I overview these main stages of the Universe's evolution.

1.1.1 The Early Universe

With Hubble's discovery of the expansion space-time (Hubble and Humason, 1931) came the realisation that the Universe started from an infinitely dense and hot point (the Big

[6] https://www.skatelescope.org/

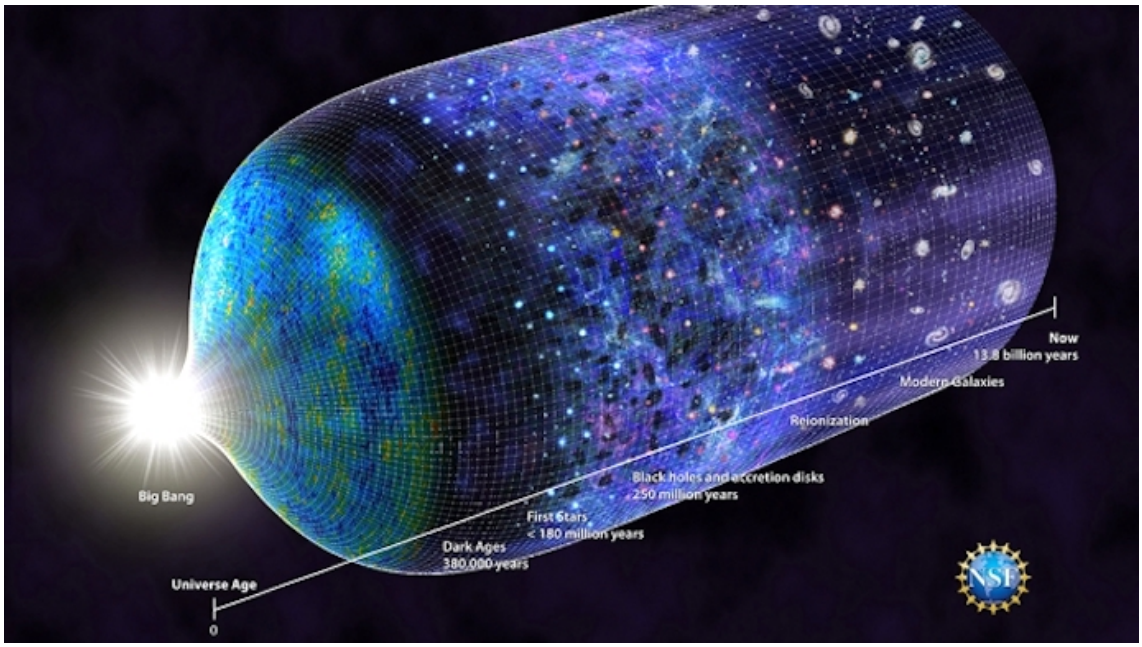

Figure 1.1: This image from the National Science Foundation shows Universe's evolution. The epoch primary considered in this work, the Cosmic Dawn, spans from the formation of the first stars to the onset of the EoR.

Bang). The currently known laws of physics are unable to describe what may have occurred when the Universe was younger than the Planck time, 10^{-43} seconds. Initially, the Universe may have undergone an exponential expansion phase, driven by a process known as inflation (Guth, 1981). Before inflation, the minuscule size of the Universe meant that quantum fluctuations due to the uncertainty principle prevented the Universe from being fully homogeneous. During inflation, space-time expanded so rapidly that these fluctuations were magnified and imprinted onto larger scales. These density inhomogeneities were the seeds for gravitational collapse, leading to the formation of all structures we currently see around us today. Dark matter is not expected to interact electromagnetically with photons so was able to begin collapsing into these gravitational wells and form halos.

Baryonic matter initially consisted of quarks and leptons as the temperature of the Universe was too high to allow them to combine into hadrons. After 10^{-6} seconds the Universe had cooled sufficiently to allow hadrons to form and about 1 second after the Big Bang singularity leptons were able to form. Once the temperature had fallen such that particle-antiparticle pairs could no longer be formed the majority of the baryons annihilated, leaving a small residue of matter for reasons we do not yet fully understand, and leading to the photon epoch. Big Bang nucleosynthesis occurred after this, leading to the produc-

tion of hydrogen nuclei, along with most of the Universe's helium-4 and trace amounts of deuterium, helium-3 and lithium-7 (Gamow and Teller, 1939). The resulting plasma was still coupled to radiation through Compton scattering of the photons and so was unable to follow the dark matter in collapsing into halos.

1.1.2 Recombination

This inappropriately named global transition actually refers to the first time protons and electrons combine to form hydrogen atoms. The expansion of the Universe eventually cooled the matter-radiation fluid sufficiently that electrons and protons had energies less than the hydrogen ionisation energy (13.6eV). After the hydrogen atoms form the Universe goes from being optically thick (photons have a very short mean free path to electron scattering, i.e. opaque) to optically thin (photons have a very long mean free path, i.e. transparent) and photons are then able to travel freely through the Universe. These photons make up the Cosmic Microwave Background (CMB) we observe today and provide us with our earliest observation of our Universe.

1.1.3 The Dark ages

The epoch between recombination and the formation of the first sources is known as the Dark Ages. The majority of the baryonic matter existed in the form of atoms at this stage, with a low level of ionization present due to photons from the high energy tail of the CMB blackbody distribution. The presence of neutral hydrogen means that this is the first time when 21-cm radiation is emitted. As the astrophysics of this time is relatively simple, observations of this Epoch would be immensely useful for constraining the matter power spectra (Tozzi et al., 2000; Iliev et al., 2002). However, measurements of the 21-cm signal from the dark ages are still well beyond our technological reach.

As the baryonic matter and radiation were decoupled the baryons were able to begin falling into existing dark matter halos. Eventually, when enough baryonic matter had fallen into these halos, star formation began. The Dark Ages ended with the formation of these first luminous sources (Schneider et al., 2002) and the Universe moved into its final global phase transition: reionization.

1.1.4 The Epoch of Reionization

The EoR spans from the formation of the first sources to almost complete ionization of the IGM. With the collapse of baryons into the dark matter halos during the dark ages star formation could commence. The radiation from these first stars (Pop III stars), their remnants and their descendants (Pop II stars) are believed to have ionized the IGM until it reached the almost fully ionized state we observe around us today. Unlike recombination, this process was not a rapid transition but an extended, inhomogeneous process with ionized bubbles growing around the sources and eventually merging together (Madau et al., 1997; Iliev et al., 2006; Mellema et al., 2006b).

The hydrogen is expected to have produced a large amount of 21-cm radiation via the hyperfine spin-flip transition. This radiation has now redshifted to the radio band and will provide us with our more promising observational probe: the 21-cm signal.

The Cosmic Dawn

In the very first stages of the EoR, before the Universe was significantly ionized, X-ray radiation from these sources was capable of heating the IGM (Madau et al., 1997). The signal is dominated by fluctuations in temperature and the Lyman-α background (details on this are given in section 1.3). We can obtain information from the CD about the X-ray spectra as these higher energy photons drive the long-range heating of the IGM. The signal also encompasses information about the radio spectra of these sources as Lyman-α photons produced are responsible for decoupling the 21-cm signal from equilibrium with the CMB (Wouthuysen, 1952a; Field, 1958, 1959; Madau et al., 1997).

Very little is known about the sources present in the CD as the formation, evolution and death of stars in primordial gas is not well understood (Bromm, 2017). These mass of these stars, known as population III stars (Pop III stars), is highly uncertain. High mass stars are predicted to be short-lived, so the 21-cm signal from the CD may be essential to our understanding of them as it is likely they will have died before reionization is underway.

Later stages

The signal from later in the EoR will be able to provide us with maps of the distribution of neutral hydrogen, with 'holes' where ionization has occurred. From statistical properties

of this signal, and eventually the size and distribution of the bubbles, we can begin to build up a picture of the evolution of luminous sources.

For more information on the EoR and Cosmic Dawn, we refer the interested reader to Furlanetto et al. (2006).

1.2 The Nature of the First Sources

As there is a lack of direct observations, the properties of the sources powering the EoR are largely uncertain. While more exotic luminous sources have been proposed, for example, dark matter annihilation (e.g. Belikov and Hooper, 2009; Liu et al., 2016), in our works we consider it prudent to focus on the high redshift counterparts of the luminous sources we see around us today.

1.2.1 Halos and Minihalos

Before recombination the first gravitationally bound structures consisted purely of dark matter as baryonic matter was coupled to radiation. As described above, halos form from density perturbations arising from the quantum fluctuations present before inflation. These fluctuations must have entered the horizon scale (i.e. the distance which light has travelled since inflation) before they can begin collapsing. In addition, these fluctuations must also be above a certain mass in order to collapse, or the random motion of the dark matter fluid causes the structure to dissipate.

In the linear regime this minimum mass able to undergo gravitational collapse is given by the filtering mass, $M_{\rm F}$, which can be thought of as a cosmological analogue for the Jeans mass. In a ΛCDM universe this is given by:

$$M_{\rm F} = \frac{4\pi\overline{\rho}_{\rm m}}{3}\left(\frac{\pi a}{k_{\rm F}}\right)^3, \tag{1.1}$$

where $\overline{\rho}_{\rm m}$ is the average matter density and $k_{\rm F}$ is the wavenumber of the density perturbation. If dark matter were hot (i.e. particles travelled with ultra-relativistic velocities) this mass would be far greater. The power spectra of cold dark matter was shown to have the statistical properties required to produce the large scale clustering of galaxies by Peebles (1982). This was shown explicitly with N-body simulations a few years later by Davis et al. (1985) and ΛCDM universe was the model favoured thereafter.

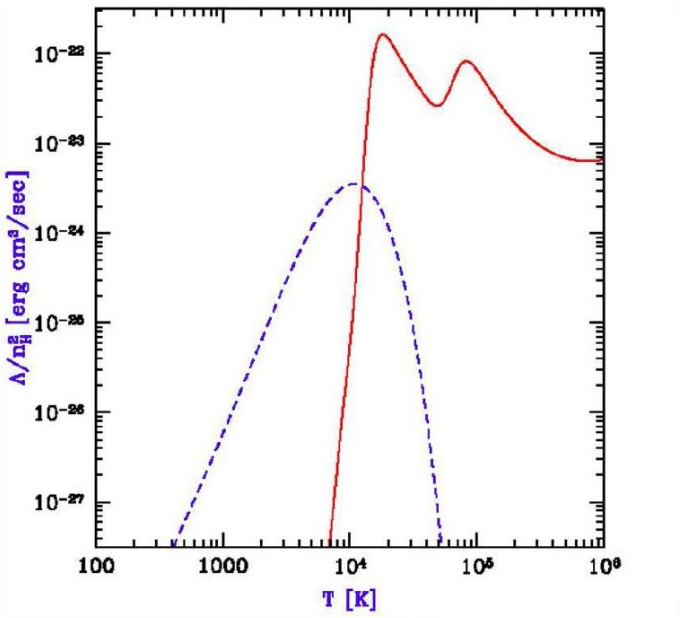

Figure 1.2: The cooling rate of primordial gas as a function of temperature. The solid red line corresponds to atomic hydrogen line cooling and the dashed blue line to molecular hydrogen cooling. Above temperatures of 10^4K the gas is cooled efficiently by hydrogen atoms, however, at temperatures below this cooling relies on the presence of hydrogen molecules. Plot from Bromm (2013b).

Once the density fluctuation is large enough non-linear gravitational collapse begins and structures slightly smaller than the filtering mass can be formed. These non-linear objects become gravitationally bound and begin to merge together, combining their masses to create larger and larger structures, a process known as hierarchical structure formation (Peebles and Dicke, 1968; Peebles, 1970, 1973, 1974, 1980).

In order to become star-forming, halos must be able to accrete baryons. To become trapped in the gravitational well created by the dark matter halo, baryons must have an average kinetic energy of less than half of their gravitational potential energy (virial theorem). Therefore, in order to accrete a gas of ions mass $m_{\rm p}$ onto a dark matter halos of mass M baryonic gas must have a temperature:

$$T_{max} < \frac{GMm_{\rm p}}{3k_{\rm b}R}, \tag{1.2}$$

where $k_{\rm b}$ is the Boltzman constant and R is the radius of the halo. Gas above this temperature is not gravitationally bound and can escape the halo.

The adiabatic temperature of the universe is well below T_{max} of many halos, however, as gas collapses into halos it is shock heated, raising the kinetic temperature. Therefore, in

order for the gas to remain in the halo it must be cooled. The dominant cooling mechanism at the temperatures of interest of primordial gas is radiative cooling through bound-bound transitions. Hydrogen is the dominant element at this time, and atomic hydrogen line cooling can bring the gas to 10^4K (see the red line in Figure 1.2). This enables halos of $\sim 10^8 M_\odot$ and above to accrete gas, and hence become star forming galaxies.

Molecular hydrogen is capable of reducing the temperature further. We can see from Figure 1.2 that for sufficiently massive halos molecular hydrogen is able to cool the gas to around 400K. Therefore if hydrogen molecules are present then the minimum star-forming halo size is smaller (Bromm et al., 1999). These smallest of halos are known as minihalos and are thought to have masses around $10^5 - 10^8 M_\odot$. However, molecular hydrogen can be dissociated by Lyman-Werner radiation, which is produced in large amounts by stars. Once hydrogen molecules are no longer present, the gas can no longer cool and so is able to escape the halo when it is heated, resulting in the suppression of star formation in the minihalo (Shapiro et al., 2004). In fact, as soon as star formation occurs it is very likely that the minihalos heat and suppress themselves meaning that star formation is short-lived. As the global star formation rate (SFR) increases a Lyman-Werner background is build up, and minihalos are unable to form due to the lack of molecular hydrogen. Minihalos are thought to have existed long before reionization completes although they may have contributed to the process early on (e.g. Haiman et al., 2000; Ahn et al., 2009, 2012).

1.2.2 Stars

Little is known about primordial star formation and the still elusive Population III (Pop III) that resulted from it. The forthcoming radio interferometer experiments are apt to place observational constraints on these sources. Pop III stars may also be glimpsed for the first time by the upcoming James Webb Space Telescope[7] (JWST) or the planned generation of ground-based giants, for example, the Giant Magellan Telescope[8] (GMT), the Thirty-Meter Telescope[9] (TMT) and the European Extremely-Large Telescope[10] (E-ELT).

These stars are thought to have been much larger than their lower redshift counterparts,

[7] https://www.jwst.nasa.gov/
[8] https://www.gmto.org/
[9] http://www.tmt.org/
[10] http://www.eelt.org.uk/

with predicted masses as high as $300M_{\odot}$ (Ahn et al., 2012), where $M_{\odot}$ is the mass of the sun. Due to their size, larger stars would be able to burn through their fuel much more rapidly leading to shorter lifetimes. Once this has occurred it is theorized that these stars would then explode in violent pair-instability supernovae (Pan et al., 2012; Whalen et al., 2013a), core-collapse supernovae (Whalen et al., 2013b) or directly collapse into black holes (Madau and Rees, 2001). In the case of direct collapse black holes, it is possible that the remnants of these stars would not emit any X-ray emission as the majority of the gas would have been stripped from the halo. Therefore, the main contribution of these stars to the reionization process is through ionizing photons, Lyman-Werner and Lyman-α, the latter two of which greatly impact the CD via suppressing minihalo formation (described above) and decoupling the 21-cm signal from the CMB respectively (see section 1.3 for more details on this). In addition to this Lyman-α photons are also capable of heating the IGM (Madau et al., 1997), but likely heating saturates early, leading to little heating or even cooling: Chen and Miralda-Escudé (2004); Meiksin (2006); Rybicki (2006); Chuzhoy and Shapiro (2006, 2007). For more details on the formation of the first stars, the interested reader is referred to Bromm (2013a).

The oldest, metal-poor stars we observe locally today, known as population II or Pop II stars, are born from the IGM once it has been enriched with heavier elements from Pop III stars. These longer living stars are capable of providing ionizing photons throughout the EoR along with copious amounts of Lyman-α and Lyman-Werner photons. It is generally thought that these stars drive reionization (Bunker et al., 2007), although there has been some speculation as to the contribution of QSOs (Madau and Haardt, 2015). These stars continue to produce copious amounts of Lyman-Series photons, providing a strong Lyman-Werner background which prevents minihalo formation throughout the later stages of the EoR.

1.2.3 High Mass X-ray Binaries

High Mass X-ray Binaries (HMXBs) are systems consisting of an early-type massive star and stellar remnant (either a black hole or neutron star), where matter from the star is accreting onto its compact companion. HMXBs dominate the X-ray emission from low redshift galaxies, so are likely to provide a significant proportion of X-rays in the high-z Universe. However, local observations suggest that these objects do not contribute a significant amount of ionizing photons due to obscuration (e.g. Lutovinov et al., 2005).

HMXBs are a promising candidate for early X-ray heating. Simulations suggest that large numbers of HMXBs will be present from the early stages of the Cosmic Dawn as the first stars are thought to have been very short-lived (Glover and Brand, 2003). The X-ray luminosity of HMXBs has been found to be proportional to the SFR Grimm et al. (2003); Chandar et al. (2009); Mineo et al. (2012a, e.g.). This relation between HMXBs and the rate at which stars are formed is due to the relatively short lifetimes of massive stars (Shtykovskiy and Gilfanov, 2007). Madau and Fragos (2017) find that the SFR and X-ray emissions are greater in more distant galaxies, which also suggests HMXBs contributed more X-ray photons.

HMXBs formation in primordial gas has been hypothesized to produce more HMXBs at high than low redshift, due to more wind accretion in low metalicity environments (Linden et al., 2010). Numerous studies have looked at HMXB evolution over cosmic time (e.g. Mirabel et al., 2011; Justham and Schawinski, 2012; Fragos et al., 2013a,b; Fialkov et al., 2014). Fragos et al. (2013a) estimate the total energy output at $z \sim$ 6-8 and conclude that HMXBs may well outshine the high redshift QSO population.

Recently ultra-deep surveys have begun to probe the galaxy population responsible for reionization (see section 1.4.3 for more details). These observational probes allow constraints to be placed on the evolution of HMXBs by looking at local galaxies with low-metalicity and compact morphologies thought to be similar to high-redshift galaxies. Some observations show a larger contribution to the X-ray budge from HMXBs (e.g. Basu-Zych et al., 2013), in agreement with predictions that HMXBs are more common in low-metalicity environments, while others find the analogue galaxies to be in agreement with the standard XLF (Basu-Zych et al., 2016, e.g.).

1.2.4 QSOs

Quasi-Stellar Objects (QSOs) are supermassive black holes at the centre of galaxies with large amounts of matter accreting onto them. The first QSOs observed were 3C 48 and 3C 273 in the 1950s (Matthews and Sandage, 1963). Initially, these objects were divided into many categories but eventually were linked by the 'unification theory', that suggested they were in fact observations of the same systems but at different orientations (e.g. Begelman et al., 1984; Rees, 1984).

The seeding mechanisms of supermassive black holes (SMBH) are not well understood. Potential seeding processes include the direct collapse of gas in the galactic centre (e.g.

Begelman et al., 2006; Agarwal et al., 2012; Ardaneh et al., 2018), accretion onto the remnant of a large pop III star (e.g. Sun et al., 2017; Becerra et al., 2018), and more exotic explanations such as dissipative dark matter (e.g. D'Amico et al., 2018) and primordial black holes (e.g. Kawasaki and Tada, 2016). The large size of observed SMBHs require that their progenitors were able to form and grow more rapidly than current theoretical estimates. This rapid initial growth may be explained by super-Eddington accretion rates (Inayoshi et al., 2016; Pezzulli et al., 2017), but how SMBHs formed and grew so rapidly remains an open question. The high level of uncertainty surrounding the astrophysical processes governing QSOs makes it unclear at what redshift they began to have an impact during the EoR.

The luminosity of QSOs varies on all time-scales, in fact, the fluctuations over very short time-scales led to the conclusion that there was indeed a supermassive black hole in their centre and are still used to probe the geometry of these systems (e.g. Kaspi et al., 2000). In addition to this, the galactic nuclei are not continually active, meaning that QSOs switch on and off depending on the accretion onto the supermassive black hole. This is known as the duty cycle of the QSO. When active the mass of the central black hole does not decide its luminosity (Kaspi et al., 2000).

Like HMXBs low redshift observations suggest QSOs do not produce a significant amount of ionizing photons (Tozzi et al., 2006). It has been proposed that QSOs contribute to reionization, however, this remains unclear (Miralda-Escudé, 1993; Meiksin and Madau, 1993; Meiksin, 2005; Giallongo et al., 2015; Puchwein et al., 2015). If QSOs are indeed present during the EoR their greatest impact of the 21-cm signal is likely to be long-range heating during the CD (Pritchard and Furlanetto, 2007).

1.3 Physics of the HI 21-cm Signal

The 21-cm signal originates from the forbidden hyperfine transition between the triplet to the singlet state in the ground state of neutral hydrogen. The longevity of the triplet state makes it ideal to trace the distribution of hydrogen over cosmic time-scales Field (1958).

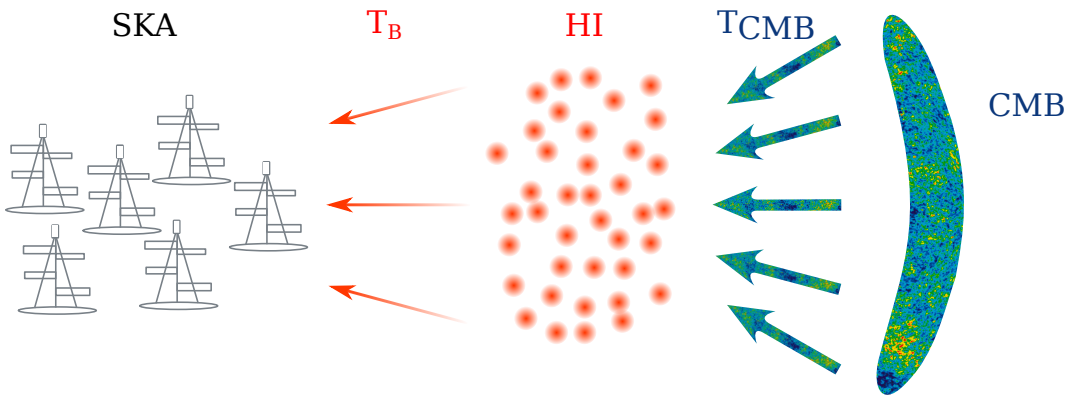

Figure 1.3: Schematic showing the 21-cm signal being observed against the CMB. SKA sees the differential brightness temperature: $\delta T_{\rm b} = T_{\rm B} - T_{\rm CMB}$.

1.3.1 Spin Temperature

The relative numbers of atoms in the triplet and singlet state can be quantified using the spin temperature, T_S. This is given by:

$$\frac{n_1}{n_2} = \frac{g_1}{g_2} e^{-T_\star/T_S}, \tag{1.3}$$

where n_1 and n_2 are the numbers of atoms in the triplet and singlet states, g_1 and g_2 are their degeneracies and $T_\star$ is the temperature corresponding to the 21-cm transition (0.0681K).

1.3.2 The Differential Brightness temperature

The brightness temperature is the temperature that a blackbody in equilibrium would have to be in order to produce an observed intensity at a specific frequency. When measuring redshifted 21-cm radiation the differential brightness temperature is used as the signal is seen with respect to the CMB. That is, we are looking at the difference between the brightness temperature of the 21-cm photons from the CMB and the brightness temperature of the 21-cm signal from the hydrogen atoms rather than an absolute temperature (i.e. $\delta T_{\rm b} = T_{\rm b} - T_{\rm CMB}$).

The signal is observed in emission ($T_S > T_{CMB}$) or absorption ($T_S < T_{CMB}$) and can be as quantified by the differential brightness temperature, δT_b:

$$\delta T_b(z) = \frac{T_S(z) - T_{CMB}(z)}{1+z} \left(1 - e^{-\tau(z)}\right), \tag{1.4}$$

where τ is the optical depth and can be written as:

$$\tau(z) = \frac{3\lambda_0^3 A_{10} T_* n_{\rm HI}(z)}{32\pi H(z) T_{\rm S}} \left(1 + \frac{1}{H(z)} \frac{dv_{||}}{/dr_{||}}\right)^{-1}, \tag{1.5}$$

where $\lambda_0 = 21.1$ cm is the line rest-frame wavelength, $A_{10} = 2.85 \times 10^{-15}\,{\rm s}^{-1}$ is the Einstein A-coefficient for spontaneous emission from the triplet to the singlet state, $n_{\rm HI}$ is the density of neutral hydrogen and $dv_{||}/dr_{||}$ is the gradient of the proper velocity along the line of sight (including both the Hubble expansion and the peculiar velocity).

The Universe is generally very transparent to 21-cm radiation, so τ is much less than one and $1 - e^{-\tau(z)} = \tau$ is true. Therefore $\delta T_{\rm b}$ can be written as:

$$\delta T_{\rm b}(z) = \frac{T_{\rm S}(z) - T_{\rm CMB}(z)}{T_{\rm S}(z)} \frac{3\lambda_0^3 A_{10} T_\star n_{\rm HI}(z)}{32\pi H(z)(1+z)} \left(1 + \frac{1}{H(z)} \frac{dv_{||}}{/dr_{||}}\right)^{-1}. \tag{1.6}$$

From this equation we can see that $\delta T_{\rm b}(z)$ can be impacted by the redshift, and the density and spin temperature of neutral hydrogen.

1.3.3 Coupling Mechanisms

In the absence of additional physics, the spin temperature is coupled to that of the CMB, as it absorbed and re-emits 21-cm photons from this source. It can be decoupled via collisions (only significant at very high redshift and in dense filaments) and the Wouthuysen-Field effect - the transition from singlet state to the triplet state via the absorption of a Lyman-α photon. This coupling is described by:

$$T_S = \frac{T_{CMB} + y_\alpha T_\alpha + y_c T_K}{1 + y_\alpha + y_c}. \tag{1.7}$$

Where T_{CMB} is the CMB temperature, T_α is the $Ly\alpha$ colour temperature ($T_\alpha = T_k$ due to numerous scatterings), y_α is the L_α coupling constant , T_k is the kinetic temperature and y_c is the collisional coupling constant.

Collisional Coupling

In dense gas, atoms can be collisionally excited or de-exited between the hyperfine levels. Collisions mainly occur between hydrogen atoms and other atoms or electrons - although there are a sub-dominant number of collisions with other species such as bare protons, deuterium atoms and helium atoms and ions (Smith, 1966). The collisional coupling coefficient is given by:

$$y_c = \frac{C_{10}}{A_{10}} \frac{T_*}{T_{\rm K}}, \tag{1.8}$$

where C_{10} is the collisional de-excitation rate from the triplet state.

The Wouthuysen-Field Effect

The Wouthuysen-Field (WF) effect, named after Wouthuysen (1952b) and Field (1958), is a less intuitive coupling mechanism. The hyperfine levels of the HI ground state can be changed indirectly via the absorption and spontaneous emission of a Lyman-α photon (or any Lyman photon). The coupling coefficient can be written as:

$$y_\alpha = \frac{P_{10}}{A_{10}} \frac{T_*}{T_\mathrm{k}} = \frac{4P_\alpha}{27A_{10}} \frac{T_*}{T_\mathrm{k}}, \tag{1.9}$$

where P_{10} radiative de-excitation rate due to Lyman-α photons and P_α is the total Lyman-α scattering rate.

1.3.4 Evolution of the 21-cm signal

In Figure 1.4 we can see the evolution of the 21-cm over time. Initially, during the Dark Ages, the signal is only impacted by the collisional coupling coefficient. This is due to the fact the Universe is sufficiently dense for collisions to occur frequently enough to decouple the signal from the CMB. Collisional coupling is effective both in the high redshift Universe (Scott and Rees, 1990) and also in dense filaments in the low redshift Universe (Kooistra et al., 2017). During this time the spin temperature can be written as:

$$T_\mathrm{S} = \frac{T_\mathrm{CMB} + y_c T_k}{1 + y_c}. \tag{1.10}$$

By the end of the Dark Ages, collisional coupling is already negligible (i.e. $y_\mathrm{c} \sim 0$). The spin temperature is uncoupled from the CMB via the WF effect. Therefore the T_S now depends on the coupling from the WF effect:

$$T_\mathrm{S} = \frac{T_\mathrm{CMB} + y_\alpha T_k}{1 + y_\alpha}. \tag{1.11}$$

As the Cosmic Dawn approaches a background of Lyman-α photons is built up to mean that the signal is no longer sensitive to fluctuations in the Lyman-α background. This is referred to as Lyman-α saturation and as $T_\mathrm{CMB} \ll y_\alpha T_k$ and $1 \ll y_\alpha$ the T_S can now be written as:

$$T_\mathrm{S} = T_\mathrm{K}. \tag{1.12}$$

It is uncertain exactly when this transition takes place due to uncertainties in the nature of Pop III stars. As the first stars are likely to produce large amounts of Lyman-α photons this assumption may be valid for the majority of the CD (Ciardi and Madau, 2003).

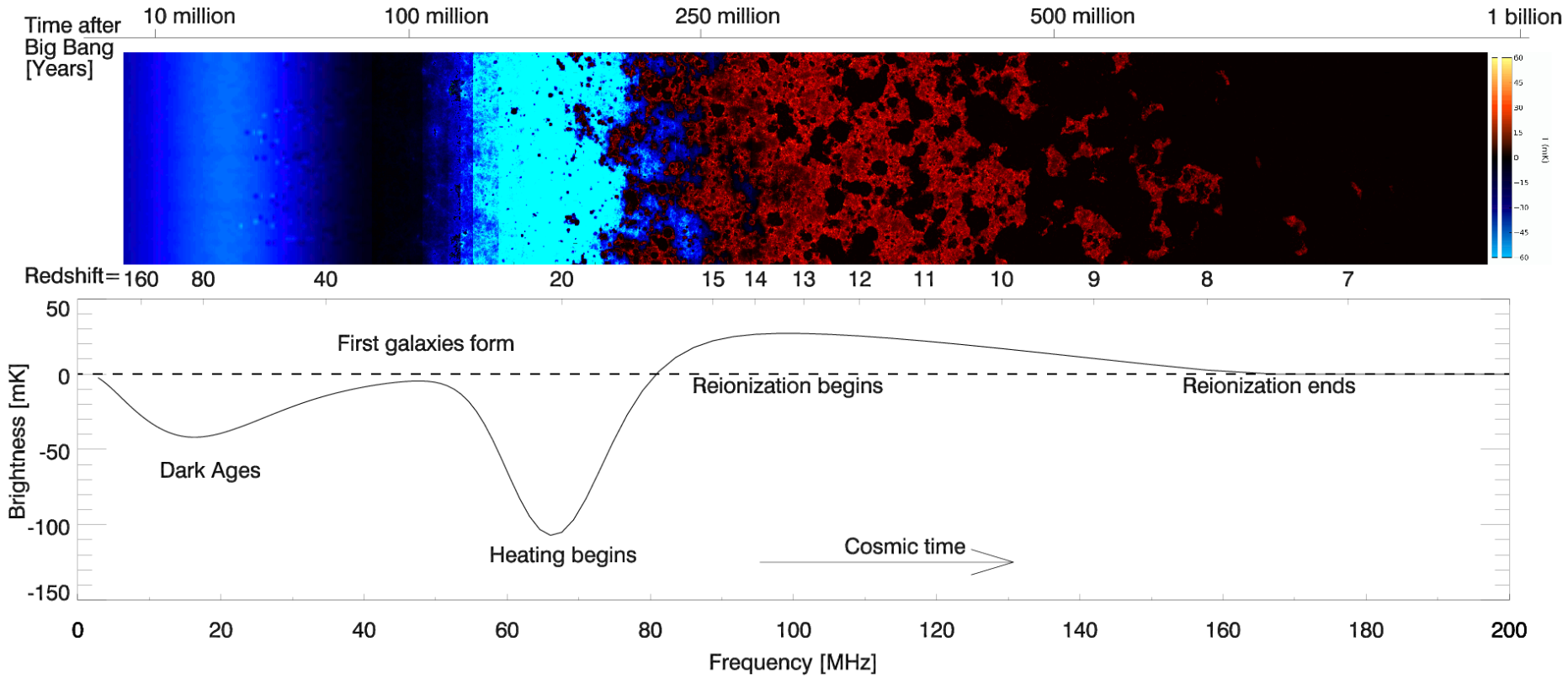

Figure 1.4: The evolution of a simulated 21-cm signal from the formation of the first sources to the end of the EoR from Pritchard and Loeb (2012). The top panel shows a lightcone of the 21-cm evolution with the colour corresponding to the magnitude of the 21-cm signal. The lower panel shows the mean subtracted 21-cm signal with key events labelled.

However, it is also possible that y_α photon density was not sufficient in certain regions to allow coupling via the WF effect through the CD (Roy et al., 2009).

As heating progresses the temperature is expected to rise above that of the CMB until $T_k \ll T_{CMB}$. Once this occurs δT_b no longer depends on the spin temperature of the gas and is only impacted by the ionized bubbles and can simply be written as:

$$\delta T_{\rm b}(z) = \frac{3\lambda_0^3 A_{10} T_\star n_{\rm HI}(z)}{32\pi H(z)(1+z)} \left(1 + \frac{1}{H(z)} \frac{dv_{||}}{/dr_{||}}\right)^{-1}. \tag{1.13}$$

Again, the time at which this occurs is uncertain due to the lack of information about sources at this redshift. Temperature saturation does not necessarily happen before reionization begins, as it is highly dependent on the X-ray sources driving the heating in the CD.

1.4 Observational Constraints

The majority of the observational constraints placed on the illusive EoR are indirect observations. However, there has recently been a possible detection of the Cosmic Dawn, which is discussed in more detail below.

1.4.1 CMB

The measurement of τ, the optical depth from the surface of the last scattering for CMB photons, constrains the time by which reionization must be underway. τ is determined by the electron density along the line of sight as their mean free path depends on the level of Thompson scattering from electrons. This dependence can be written as:

$$\tau = \int_0^{z_{\rm rec}} \frac{\sigma_{\rm T}\ n_{\rm e}\ c}{H_0(1+z)\sqrt{\Omega_{\rm m,0}(1+z)^3 + \Omega_{\Lambda,0}}}\ dz, \tag{1.14}$$

where $\sigma_{\rm T}$ is the Thompson scattering cross-section, and $n_{\rm e}$ is the electron number density.

Thompson scattering of CMB photons from free electrons alters the polarization of the CMB, and from this τ can be inferred. Using this the Planck collaboration most recently calculated the value of τ to be 0.058 ± 0.012 (Planck Collaboration et al., 2016). As this is an integrated quantity over time it cannot provide us with any information about the

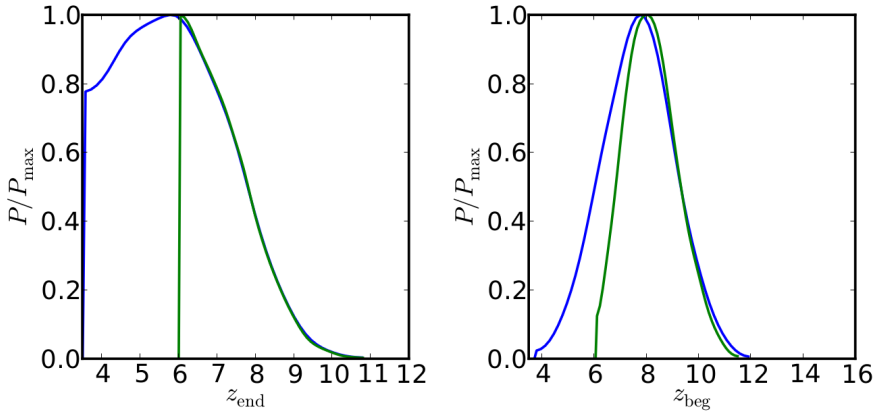

Figure 1.5: The strongest constraint for the beginning of reionization comes from the Planck collaboration. Here posterior distributions of $z_{\rm end}$ and $z_{\rm beg}$ from Planck Collaboration et al. (2016) from redshift symmetric parametrization are shown. Planck and VHL data and information from the kSZ amplitude are combined with the current data to obtain the above results. The blue and green lines show data without and with (respectively) the constraint from Lyman-α Forest measurements that $z_{\rm end} \approx 6$.

progression of reionization. In fact, when calculating the redshift at which reionization took place it is often assumed that the process is instantaneous, from which they calculate a value of $z_{\rm reion} = 8.8 \pm 0.9$. In order to make their model more realistic they also estimate the width of such a transition. In combination with the Lyman-α Forest constraints (discussed in section 1.4.2), this gives an estimated time for the beginning of reionization (i.e the point at which the ionized fraction was 10%) to be $z_{\rm beg} = 8.1^{+1.1}_{-0.9}$ (see Figure 1.5).

1.4.2 Lyman-α Forest

We can learn about the distribution of neutral hydrogen at the end of the EoR using the Lyman-α Forest. Hydrogen atoms scatter radiation from background QSOs out of the line of sight via a Lyman-α transition (n=1 $\rightarrow$ n=2) which gives rise to absorption lines in the QSO spectra known at the Lyman-α Forest. The frequency at which these absorption lines allows the redshifts were neutral hydrogen is present along the line of sight to be inferred. These spectra provide a measure of the ionized fraction of the IGM along the line of sight and can be used to detect the end of reionization. It was first predicted in 1965 that QSO spectra would exhibit an absorption trough in the presence of large amounts of

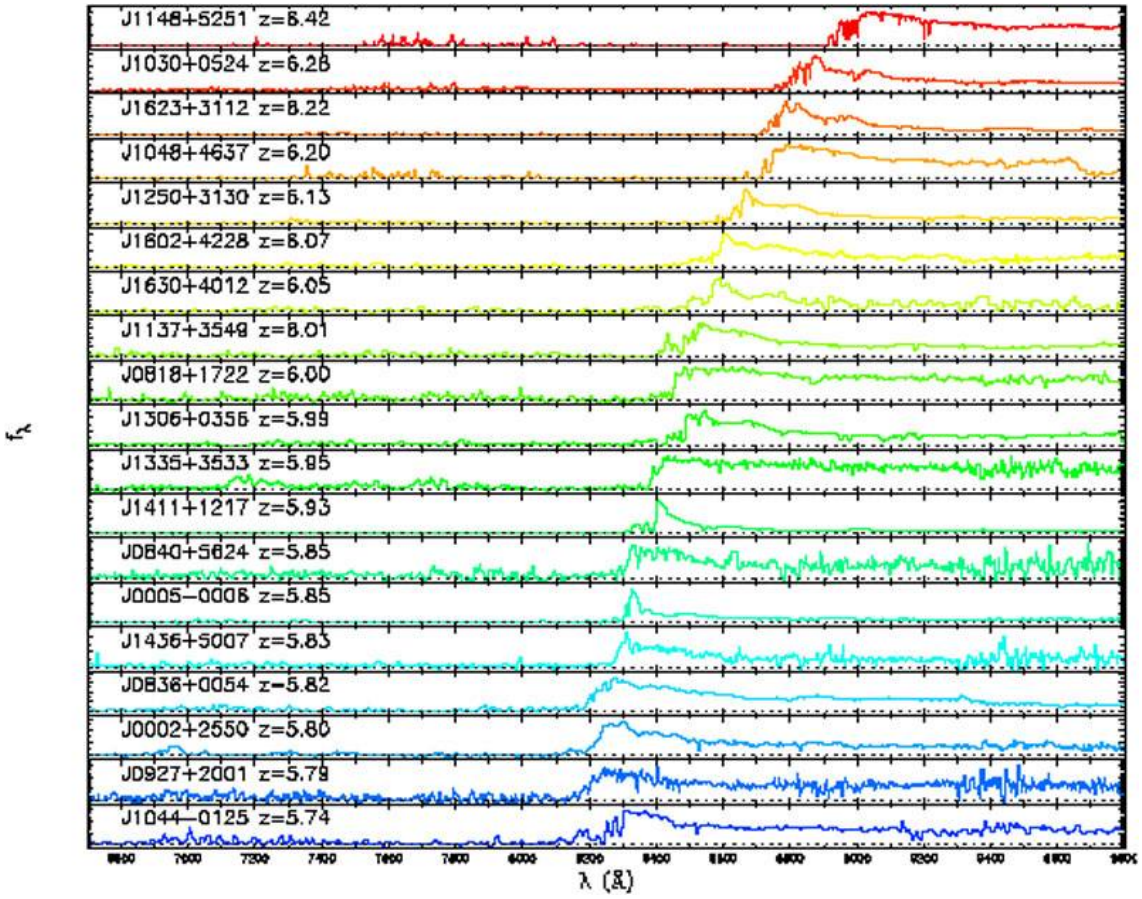

Figure 1.6: Spectra of QSOs from the Sloan Digital Sky Survey (Fan et al., 2003, 2006) from Zaroubi (2013). Here we can see that the Gunn-Peterson trough is shifted to lower frequencies as the redshift increases, which is often attributed to the increase of the ionized fraction with redshift as reionization ends. However, the quantitative increase in ionized fraction cannot be determined from these spectra.

neutral hydrogen, known as a Gunn-Peterson trough (Gunn and Peterson, 1965) and first observed in 2001 by Becker et al. (2001).

Advances in hardware, namely the addition of high-resolution echelle spectrographs to large telescopes, have led to a huge increase in the quality of the data available. In the meantime, the development of ΛCDM has led to a theoretical explanation for all the features seen in the Lyman-α Forest. The spectra of these high redshift quasars from the Sloan Digital Sky Survey (Fan et al., 2003, 2006) have shown the Gunn-Peterson trough generally occurs at a lower frequency for higher redshifts as radiation from higher redshift QSOs has to travel through more neutral hydrogen (see Figure 1.6). There are a few examples (J00005-0006 at z=5.85 and J1411+1217 at z=5.93) which have fewer absorption lines in their troughs when compared to other spectra from similar redshifts

which may point towards patchy reionization. Currently, the QSO at the highest redshift ever detected is ULAS J1120+0641 at z=7.084 (Mortlock et al., 2011).

The Lyman-α Forest has been used to calculate the optical depth along the line of sight and in doing so has constrained the end of reionization to z>$\sim$6. However, there continues to be disagreement as to how ionized the Universe was prior to this time.

Hotter atoms produce wider absorption lines due to Doppler broadening. This means the Lyman-α forest is not only sensitive to the density of neutral hydrogen but also its temperature. This information is able to constrain the reionization history as the ions have been cooling since they were ionized and photo-heated.

1.4.3 Lyman-Break Galaxies

In order to learn more about the types of sources likely to be present during reionization, we must turn to high redshift observations of galaxies. A commonly used method of observing these galaxies is the Lyman-break technique (Steidel et al., 1995). The sky is imaged with filters of increasing wavelength. Galaxies 'drop out' of the image depending on what redshift they are obscured by neutral hydrogen (as in the Gunn-Peterson trough).

While this method picks up a very large number of sources, it is possible that the breaks may correspond to another absorption line so spectroscopy is required to verify the redshift of each individual source. Large sky surveys use this technique to build up a map of high-redshift galaxies and from this estimate the global SFR and hence the total number of ionizing photons produced by high redshift galaxies. (e.g. Bouwens et al., 2011; Willott et al., 2015; Bouwens et al., 2016; Greig and Mesinger, 2017; Shivaei et al., 2018).

In Figure 1.7 the ionizing photon efficiency found from different studies of high redshift galaxies is shown. As galaxies are the most likely candidates for the sources driving reionization the global SFR can be used to estimate the total number of ionizing photons. This can loosely constrain the reionization histories providing that galaxies are the dominant producers of ionizing photons and that the correct assumptions are made about clumpiness and the escape fraction. It should be noted that the escape fraction is very poorly constrained so these constraints are not particularly accurate. Data from Lyman-Break galaxies have also been used to support reionization being a patchy rather than smooth process (Pentericci et al., 2014).

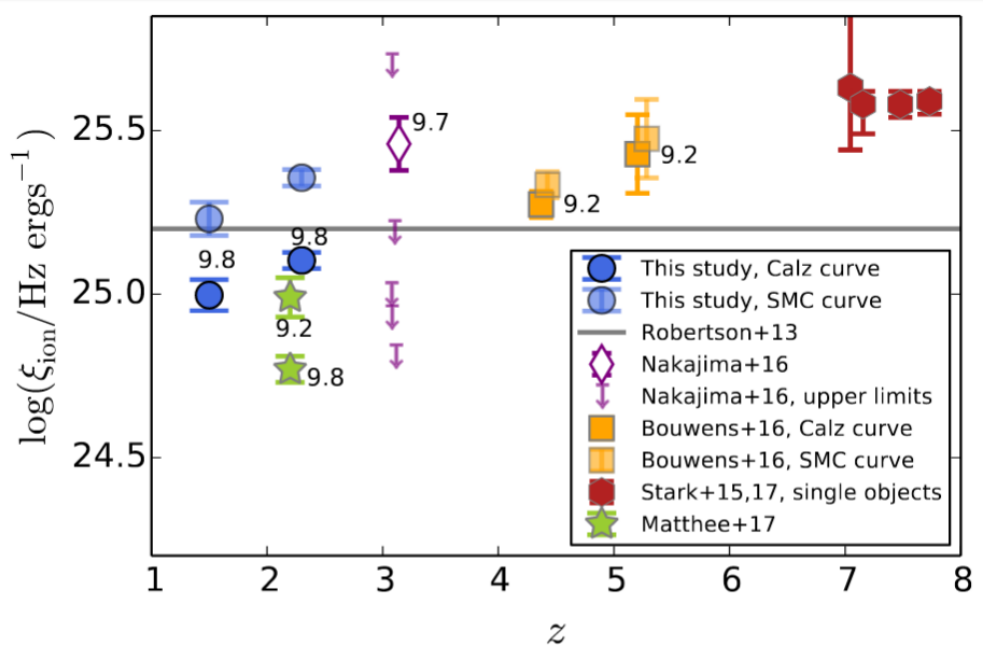

Figure 1.7: Here the ionizing photon production efficiency of high redshift galaxies found in different studies are plotted against redshift. These can be used to place a loose constraint on the reionization history of the Universe if we assume that galaxies are driving reionization. This plot was taken from Shivaei et al. (2018) and more details can be found here in the different studies.

1.4.4 The 21-cm Line

We now move to the 21-cm line, the most promising observation probe of the EoR and CD. The effects of instrument noise, galactic foregrounds and contamination from the ionosphere and radio signals human activity make this an exceptionally challenging observation. This thesis focusses on the possibility of detecting the CD with this measurement, and a detailed overview of the theory behind it will be presented in the next section (section 1.3).

LOFAR

While the LOw Frequency ARray [11], LOFAR, has not yet made a direct detection of the 21-cm signal, it has placed the first upper bounds on the power spectra of the signal (Patil et al., 2017). After foreground subtraction and removal of noise bias they claim to have detected a $\Delta 2I=(56\pm13$ mK$)2$ (1-σ) excess variance and a best 2-σ upper limit of $\Delta 22 <(79.6$ mK$)^2$ at k=0.053 h cMpc^{-1} in the range z = 9.6 - 10.6. With more data likely to become available in the near future, it is hoped the LOFAR will be able to lower this upper bound. The first models that will be either detected or excluded are late heating models, as these produce the highest power spectra, making these future observations of possible relevance to this work.

The EDGES detection

Recently the Experiment to Detect the Global EoR Signature[12], EDGES, experiment claimed to have detected an extremely strong absorption signal from the CD (Bowman et al., 2018). If this result is indeed true then additional physics is required to explain the phenomena. In the wake of the result, many ideas have been put forward. The low value of the 21-cm signal implies that either the kinetic temperature of the gas was much colder than expected or a radio background dominated the 21-cm frequency at the time. It has been hypothesized that interactions between dark matter and baryonic matter have cooled the IGM (e.g Yang, 2018; Barkana, 2018; Muñoz et al., 2018). It has also been suggested dark energy could lead to this scenario by altering the timing of reionization by accelerating the expansion of the Universe and thus leading to a lower baryonic temperature (Hill and Baxter, 2018). Deviations from the expected blackbody spectra of the CMB have been

[11] http://www.lofar.org/

[12] http://loco.lab.asu.edu/edges/

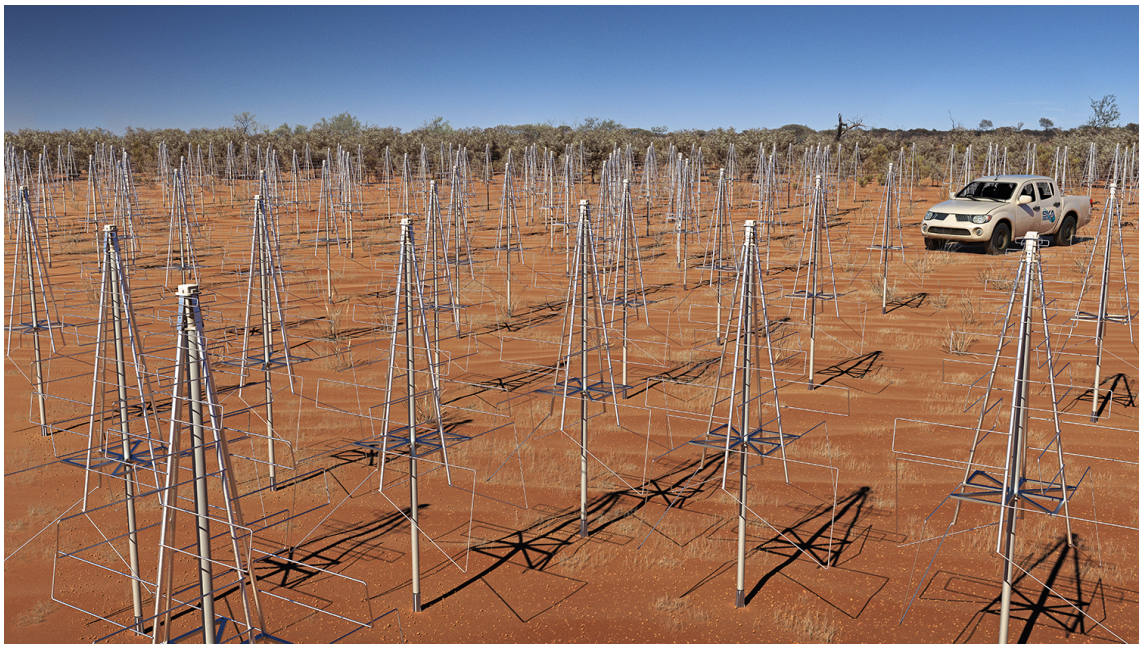

Figure 1.8: The expected appearance of SKA-LOW once it has been constructed in Western Australia. Image from www.skatelescope.org.

proposed as the source of additional 21-cm photons at the time (Pospelov et al., 2018) along with contributions from early black holes (Ewall-Wice et al., 2018). In light of the observational difficulties of this experiment and concerns over the foreground modelling (Hills et al., 2018), further validation is required by another, independent observation in order to verify the result. We also consider this scenario to be unlikely due to the extremely low temperature and the flat evolution of their 21-cm signal in the trough.

SKA

The Square Kilometre Array (SKA) is a highly anticipated future radio interferometer. SKA will be located in two separate locations; SKA-LOW in Australia and SKA-MID in South Africa. Construction is set to begin by the end of 2019. The work presented in this thesis has been completed with the view of making predictions for SKA-LOW.

SKA-LOW will consist of 512 radio antennae ($\sim$130,000 baselines) which will cover an area of 0.4 km^2, which will make it the world's largest radio interferometer. The maximum baseline will be 65 km and it will be capable of observing in the 50-130 MHz range. When compared to the current leading radio interferometer (LOFAR) it will have eight times the sensitivity, 25% better resolution and 135 times the survey speed. See the SKA website for more information and updates: www.skatelescope.org. As SKA is expected to operate

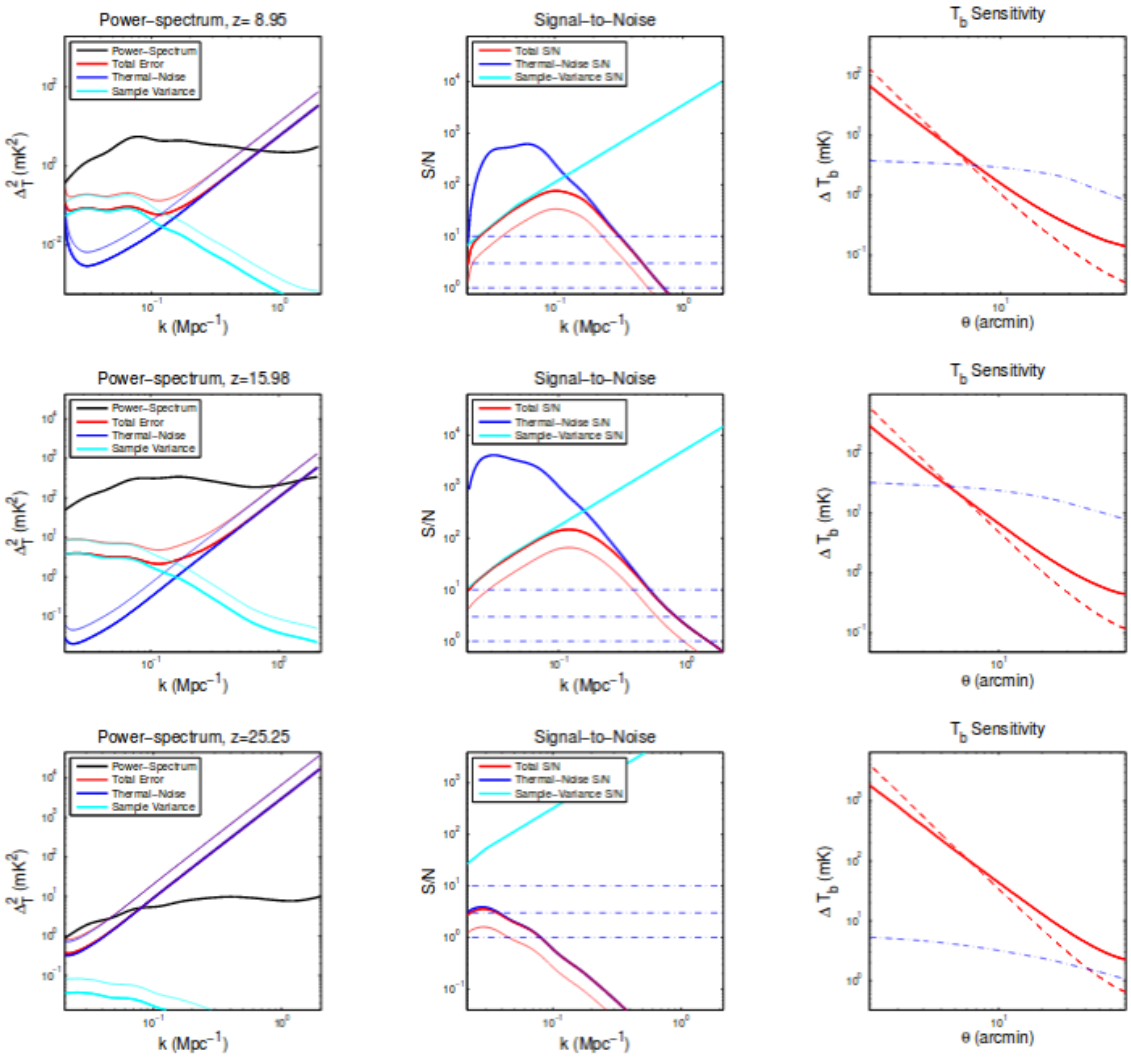

Figure 1.9: Plot from Koopmans et al. (2015) showing the expected noise for deep integrations of SKA1-Low for several redshifts. The left-hand side panels show the expected brightness temperature power-spectra from the CD 21-cm signal Mesinger et al. (2013) and the expected thermal-noise and cosmic-variance errors. The central column show the signal-to-noise ratio and the right-hand side panels show the sensitivity on $\delta T_{\rm b}$ as a function of resolution. For details on the parameters refer to Koopmans et al. (2015).

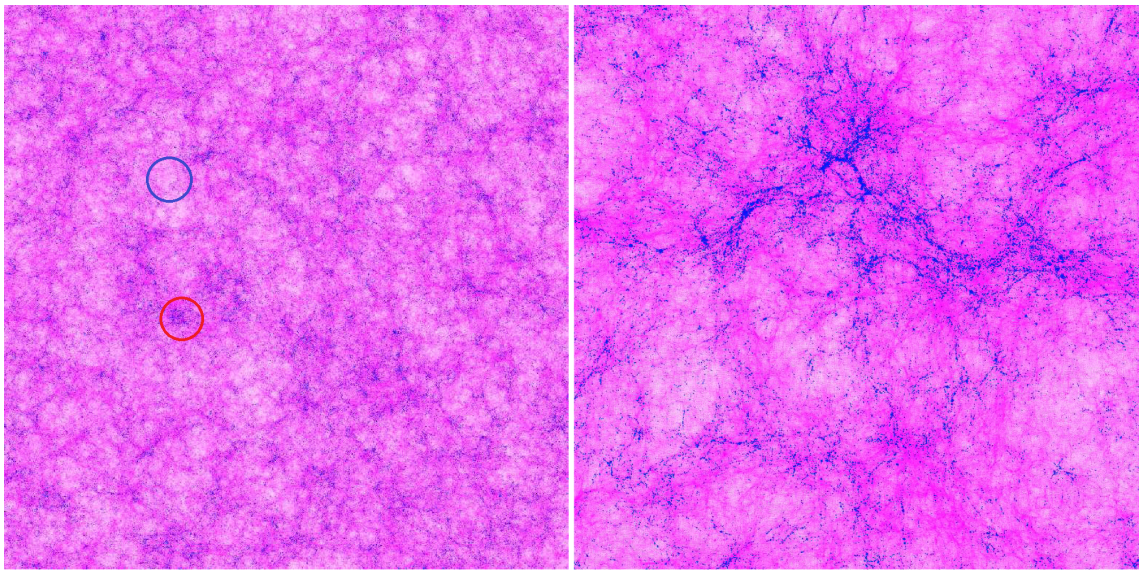

Figure 1.10: Here the density field (purple) and halos (blue) from a high resolution N-body simulation are shown. On the left the full box is shown (425 × 607 Mpc per side) and on the right a zoomed region (41.5 × 41.5 Mpc). If the simulation was run over the smaller volume more structures would be resolved, however, the large scale voids (example circled in blue) and over-dense regions (example circled in red) present in the larger box would not be included. Plot from Iliev et al. (2014).

with such unprecedented sensitivities it will be capable of detecting the CD. Using the 21-cm signal predicted by Mesinger et al. (2013), Koopmans et al. (2015) predict that it maybe be possible to image the CD on scales larger than half a degree. At lower redshifts in the EoR this improves to scales of around 5 arcmin These conclusions are, however, dependant on the heating model considered.

As it is a radio interferometer, SKA will measure the mean subtracted signal rather than an absolute value. In the later stages of reionization the mean value can be extracted using the information that the ionized bubbles give no signal. However, at the CD, this is not possible so instead of considering the absolute value produced it is prudent to instead investigate the fluctuations of the signal.

1.5 Simulating the EoR

Due to the multifarious interplaying astrophysical processes that impact the HI 21-cm signal, fully numerical simulations are a pertinent approach in the study of this Epoch. This is computationally challenging due to the multi-scale nature of the process. It is

essential to take into account the largest scale density fluctuations in order to include the rarer, large halos and astrophysical objects within them as well as voids. At the same time, the astrophysical processes driving the EoR and CD occur on kiloparsec scales. In order to simultaneously model these scales a large box with high resolution is required, making this problem extremely computationally intensive. In order to make it feasible it is necessary to trade accuracy at either large or small scales.

In Iliev et al. (2014) it is shown that a boxsize of at least 100 Mpc is required to capture reionization patchiness. However, due to the long mean free path of X-rays driving the inhomogeneous heating this is unlikely to be sufficient for the CD. As discussed above it is necessary to include sources down to 10^5 $M_\odot$ if minihalos are included. Subgrid modelling can be used to capture the smaller scale physics.

1.6 Outline of the Thesis

The outline of this Thesis is as follows: I describe the simulations, models and techniques used in this work in Chapter 2. In Chapter 3 I present the work done on the impact of X-ray heating done by X-ray binaries, which has been previously published in Ross et al. (2017). Chapter 4 focuses on simulations I have run on the impact of extreme QSO heating and the possible impact of the timing of Lyman-α models on this result. Some of this work has been previously published in (Ross et al., 2018). Finally, in Chapter 5 I present the conclusions drawn from this work. Additional information about updates to the code I have contributed to is available in the Appendices.

Chapter 2

Methodology

2.1 N-body Simulations: cubep^{3}m

N-body simulation are widely used in astrophysics to understand dynamical systems from clusters of stars to the structure formation of dark matter. Cosmological simulations model dark matter as particles that only interact with each other via gravity and have successfully matched many observables (e.g. Springel et al., 2005; Boylan-Kolchin et al., 2009). However, pure N-body simulations do not take into account the impact of the reionization process on star formation (e.g. Shapiro et al., 2004; Dawoodbhoy et al., 2017). In order to accurately quantify these baryonic effects hydrodynamical simulations must be run in order to correctly simulate them (Sullivan and Iliev, 2016). These baryonic effects only occur on small scales, so are impossible to resolve when considering large box sizes. For this reason we use N-body simulations are include some of these effects in the form of sub-grid models (see section 2.2 for more details).

Throughout this work we obtain the underlying cosmological structures and dark matter halo catalogues from the N-body simulation code CUBEP^{3}M. CUBEP^{3}M is an open source, high-performance N-body code that is particularly suited to simulate dark matter over large scales. (See Harnois-Déraps et al. (2013) for details on this code and (Dixon et al., 2016a) for more details on the exact runs used in this work.) In CUBEP^{3}M a two level particle mesh is used to calculate long-range gravitational forces which are then kernel matched to short-range (less than four times the maximum distance between particles) particle-particle interactions. The spherical over-density algorithm (integrated into CUBEP^{3}M) was used to find dark matter halos with an over-density parameter of 178. The

box contained 4000^3 particles with 349 Mpc per side volume which ensured reliable halo identification down to $10^9 M_\odot$, which is the minimum mass at which ionized hydrogen can be accreted.

2.2 Source Models

All luminous sources form within dark matter halos. Dark matter halos above $10^9 M_\odot$, known as High Mass Atomically Cooling Halos, HMACHs, are are found as described above in section 2.1 as they can be reliably identified having 40 particles each. However, halos between 10^8 and $10^9 M_\odot$ are capable massive enough to accrete neutral gas so must also be considered (at least until they are ionized, at which point they are suppressed as they are incapable of accreting ionized gas). These halos are added using the subgrid model detailed in Ahn et al. (2009). These models are calibrated to very high resolution simulations to add the haloes between $10^8 - 10^9\,\mathrm{M}_\odot$, hereafter referred to as Low Mass Atomically Cooling Halos, LMACHs.

2.2.1 Stellar Sources

Stellar sources are assumed to have black body spectra with effective temperatures of 5×10^4K, which is consistent with the observed spectra of the massive O and B stars (see Figure 2.1). The luminosity of these sources is proportional to the mass of the host halo, more details of this parametrization can be found in Chapter 3. As can be seen from the left hand side panel of Figure 2.3 these stellar sources do not emit X-ray radiation. Hence, they do not contribute significantly to the long-range heating of the IGM. However, it is important to include them in order to track the growth of ionized regions.

For a source with halo mass, M, and lifetime, t_s, we assign a stellar ionizing photon emissivity according to

$$\dot{N}_\gamma = g_\gamma \frac{M \Omega_\mathrm{b}}{\mu m_\mathrm{p} (10\,\mathrm{Myr}) \Omega_0}, \tag{2.1}$$

where the proportionality coefficient g_γ reflects the ionizing photon production efficiency of the stars per stellar atom, N_i, the star formation efficiency, $f_\star$, and the escape fraction, f_esc:

$$g_\gamma = f_\star f_\mathrm{esc} N_\mathrm{i} \left(\frac{10\,\mathrm{Myr}}{t_\mathrm{s}} \right). \tag{2.2}$$

(Haiman and Holder, 2003; Iliev et al., 2012).

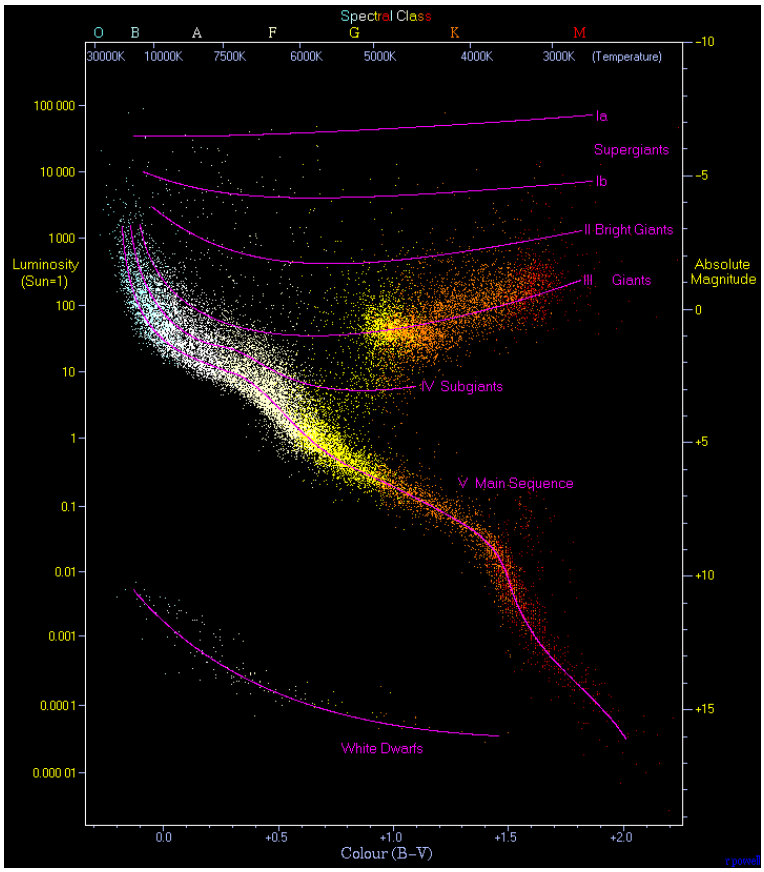

Figure 2.1: A Hertzsprung-Russell Diagram from the Hipparcos Catalogue (Perryman et al., 1997). Here we can see that the effective temperature of O-B stars is approximately 2×10^4K.

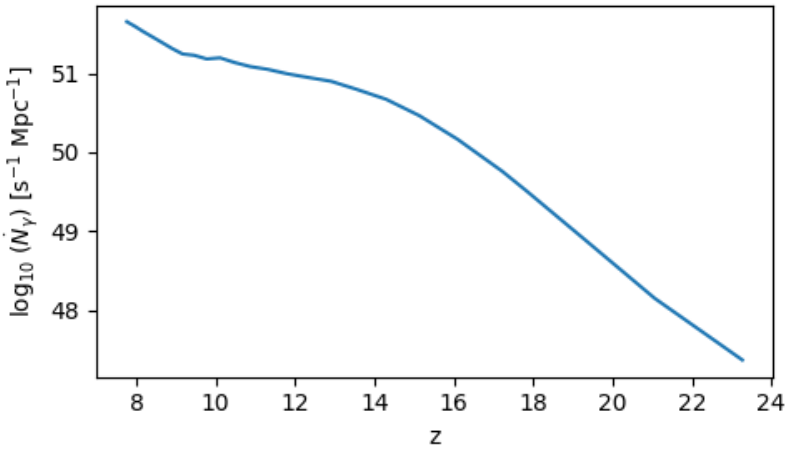

Figure 2.2: The emission rate of ionizing photons per comoving Mpc, $\dot{N}_\gamma$, is plotted against redshift. $\dot{N}_\gamma$ is calculated from simulation outputs so takes source suppression from photo-heating into account.

HMACHs have an efficiency $g_\gamma = 1.7$ and are assumed to be unaffected by radiative feedback, as their halo mass is above the Jeans mass for ionized ($\sim 10^4$K) gas. LMACHs have a higher efficiency factor $g_\gamma = 7.1$, reflecting the likely presence of more efficient Pop. III stars or higher escape fractions (Iliev et al., 2007a).

LMACHs are susceptible to suppression from photo-ionization heating. In this work, we assume that all low-mass sources residing in ionized cells (with an ionized fraction greater than 10 per cent) are fully suppressed, i.e., they produce no ionizing photons (Iliev et al., 2007a; Dixon et al., 2016a). The source model for the stellar radiation is identical to the one in simulation LB2 in Dixon et al. (2016a) and uses the aggressive suppression model 'S' defined there. However, the details of this suppression are not very significant here, since we focus on the very early stages of reionization before significant ionization develops. The stellar sources are assigned a blackbody spectrum with an effective temperature of $T_{\rm eff} = 5 \times 10^4 K$.

$f_{\rm esc}$ and f_* are still highly uncertain parameters, but our values of g_γ are within the range permitted observations. For example, if we assume $f_{\rm esc} \approx 10$ per cent (as found in Marchi et al., 2017) and $N_{\rm i}$= 50000, 6000 for LMACHs and HMACHs respectively (in agreement with Leitherer et al., 1999; Schaerer, 2002) then for our timestep of 11.5 Myrs this gives a value of $f_* \approx 0.02$ per cent, 0.03 per cent for LMACHs and HMACHs respectively. This value is in agreement with some works (for example Lilly et al. (2013) find a value of $\sim$ 0.01-0.07 per cent).

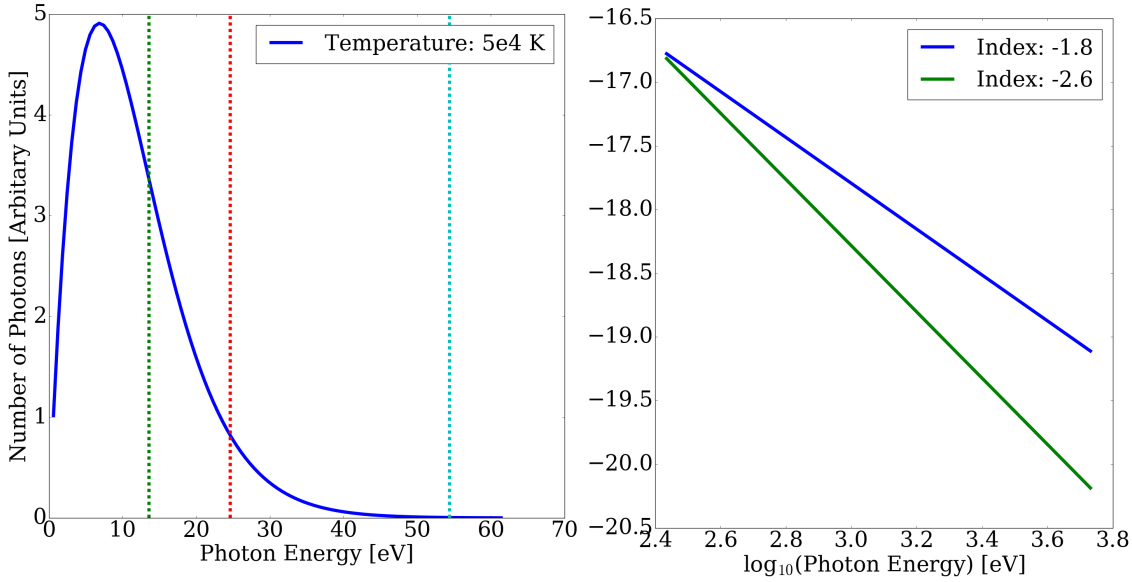

Figure 2.3: An HR-Diagram from the Hipparcos Catalogue (Perryman et al., 1997). Here we can see that the effective temperature of O-B stars is approximately 2×10^4K.

Madau and Dickinson (2014) find a slightly lower value for g_γ than used in this work (g_γ between 0.4 and 1.2 at z$\approx$ 10, see the left-hand side panel of Figure 16 in their paper). This is reasonably close to our value for HMACHs, and much lower than LMACHs and could be explained by the prescence of more luminous Pop III stars/ higher escape fractions at higher redshifts, as mentioned above (Iliev et al., 2007a). In Figure 2.2 $\dot{N}_\gamma$ is plotted against redshift. The values are a little higher than that found in Bolton and Haehnelt (2007), see Figure 7 in their paper, which may be due to pop I stars beginning to dominate over pop II stars by these later redshifts.

The Lyman-α component of the stellar spectra is also important to the 21-cm signal due to the WF effect. An SED is chosen as a realistic approximation for this region of the spectrum. The SED generated by population synthesis code PEGASE2 (Fioc and Rocca-Volmerange, 1999) were used as the spectra for Pop II stars. The metalicity was assumed to be 10^{-3}Z$_\odot$ (where Z$_\odot$ is solar metallicity) and a a Salpeter initial mass function was used, with stars with masses between 1 and 100 M$_\odot$. The luminosity was set to be consistent with the rest of the spectra.

2.2.2 High Mass X-ray Binaries

As HMXBs are formed from stars and their remnants, we assume that their luminosity is proportional to the SFR, as is found in Mineo et al. (2012a). There may be a contribution

from hot diffuse gas, but this is found to be sub dominant in Mineo et al. (2014) with diffuse emission only accounting for ≈1/3 of the X-ray emission with at about ∼3040 per cent of this attributed to faint compact sources of various types (Mineo et al., 2012b).

We take the SFR to be proportional to the stellar mass (this may well not be the case, see Mineo et al. (2012b), but we leave more realistic stellar models to future work). Therefore, HMXBs are assumed to be present wherever stars are present, with the totally luminosity of HMXBs in that galaxy proportional to the mass of the host halo. The spectra of HMXBs is a power law of with an index of $\alpha = -2.5$, see the right hand side panel of Figure 2.3. More details of these sources can be found in Chapter 3.

HMXBs sources are also assumed to reside in dark matter haloes. They are assigned a power-law spectrum with an index of $\alpha = -1.5$ in luminosity. Hickox and Markevitch (2007) showed, using the *Chandra Deep Fields*, that X-ray sources with a single power law spectrum would over-contribute to the observed X-ray background (XRB) if $\alpha \lesssim 1$, thus requiring softer spectra for the reionization sources.

The frequency range emitted by these sources extends from 272.08 eV to 100 times the ionization threshold of for doubly ionized helium (5441.60 eV) (Mineo et al. (2012a); Mesinger et al. (2013)).

The X-ray luminosity is also set to be proportional to the halo mass, since HMXBs are formed from binary systems of stars and stellar remnants. Unlike their stellar counterparts, these sources have the same efficiency factor for all active sources. Low-mass haloes that are suppressed are assumed not to produce X-ray radiation. The efficiency is parametrised as follows:

$$g_{\rm x} = N_{\rm x} f_{\star} \left(\frac{10 \text{ Myr}}{t_{\rm s}} \right) \tag{2.3}$$

where $N_{\rm x}$ is the number of X-ray photons per stellar baryon. A value of $N_{\rm x} = 0.2$ is roughly consistent with measurements between 0.5–8 keV for X-ray binaries in local, star-bursting galaxies (Mineo et al., 2012a) although the uncertainty is a factor 2 to 3 (see the discussion in Mesinger et al., 2013). We take $g_{\rm x} = 8.6 \times 10^{-2}$, which implies $f_{\star} \approx 0.4$ if $N_{\rm x} = 0.2$. Our X-ray luminosities are, therefore, somewhat higher than in the local Universe. The total number of X-ray photons contributed from these sources over the simulation time is an order of magnitude lower than the value obtained from the Cnandra Deep Fields for the XRB between 1 and 2 keV (Hickox and Markevitch, 2007) making these sources consistent with observations. Note that the long-range X-ray heating we examine here is dependent on the abundance, clustering and spectra of the X-ray sources.

2.2.3 QSOs

Due to an insufficient number of suitable redshift observations, we must make some assumptions about the luminosity and spectra of the QSOs present during the CD. We assume that our QSO-like sources only produce X-rays, which physically means that stellar sources dominate the lower frequency photon budget. (Note that by considering only the X-ray photons, we are neglecting the UV contribution from QSOs. Within our model and at these high redshifts, we do not expect a large impact on our results, but some comparisons to other studies may complicated by this fact.)

The X-ray emissivity from QSOs is quantified using the QSO X-ray Luminosity Function, QXLF. Here, we follow Ueda et al. (2014, see Section 6.2 for the functional form and Table 4 for the parameter values), though modified to reflect the uncertainty in the higher redshift behaviour. In particular, we alter the high-redshift (in this case, $z > 3$) density evolution parameter (known as $p3 = -6.2$ in Ueda et al. 2014) to a smaller -2, which is more in line with Giallongo et al. (2015) or generally including more QSOs at high redshift. Furthermore, one aim of this paper is to investigate the maximal impact of these type of sources; though, we concede that such a shallow QSO density evolution is unlikely at the highest redshifts. This QXLF takes the form of a double power law with luminosity-dependent density evolution and is taken over an X-ray luminosity ($L_{\rm X}$) range of $10^{42} - 10^{47}$ ergs s^{-1}).

The number density of QSOs in our simulation volume, $n_{\rm q}$, is calculated by integrating the QXLF, $\Phi(L, z)$, for each redshift

$$n_{\rm q}(z) = \int_{L_{\rm min}}^{L_{\rm max}} \Phi(L, z)\ dL, \tag{2.4}$$

where $\Phi \propto (1+z)^{-2}$ is the QXLF. At high redshift, the number density of haloes capable of hosting quasars (i.e., the HMACHs number density) is insufficient to replicate the QXLF (see below for host halo details). At this time QSOs are placed in all the existing HMACHs.

The QSO spectrum is assumed to be:

$$L_{\rm q}(\nu) \propto \nu^{-\alpha_{\rm x}^{\rm q}}, \tag{2.5}$$

where $\alpha_{\rm x}^{\rm q} = 0.8$ (Ueda et al., 2014) or 1.6 (Brightman et al., 2013) for our two QSO models. More generally, we chose a hard and soft model, where the exact spectral indices

are unimportant, for comparison's sake. The luminosity of each QSO at 2 keV is then assigned by randomly sampling the QXLF. Given this luminosity and spectral index, $\dot{N}_\gamma$ is calculated for the same frequency (energy) range as the HMXBs.

While black holes masses are correlated to the velocity dispersions of the bulges of their host galaxies (known as the Maggorian relation, see Ferrarese and Merritt (2000)) the luminosities of observed QSOs do not correlate with the mass of their host galaxy or that of the central black hole. Rather, they depend on the physics of the accretion disk (e.g. Woo and Urry, 2002; Middei et al., 2017). Consequently, in our simulations, we place the active QSOs in random HMACH haloes[1]. Initially, HMACHs are too rare to host sufficient numbers of QSOs to reproduce the (admittedly optimistic) luminosity function. At these early times, we assign a single QSO to each existing HMACH halo. Furthermore, in our simulations, we assume a QSO lifetime of 34.56 Myr, which is consistent with current estimates (e.g. Borisova et al., 2016; Khrykin et al., 2017). It has been long known that the luminosity of QSOs varies with time over all frequencies, including the X-ray range of the spectrum (e.g. Halpern, 1984; Pan et al., 1990; Jin et al., 2017), and we mimic this behaviour by assigning a new $L_{\rm q}$ every 11.52 Myr, with a value that is within an order of magnitude of the previous $L_{\rm q}$ while the QSO is active.

2.3 Radiative Transfer: C^2-Ray

For our radiative cosmological transfer simulations we use the fully numerical code C^2-Ray, **C**onservative, **C**ausal **Ray** tracing. C^2-Ray employs a short characteristics ray tracing method (Raga et al., 1999) which is explicitly photon conserving in both space and time. C^2-Ray has been used in many previous works (e.g. Iliev et al., 2006; Mellema et al., 2006b; Iliev et al., 2007b, 2014; Dixon et al., 2016a). In this section I discuss the algorithm, testing and development of this code.

The Algorithm

Radiative transfer methods often impose impractically short time-steps and small cell sizes. C^2-Ray was developed in order to bypass these restrictions. This is achieved by making two assumptions:

[1] In rare cases, an HMACH will disappear for some reason, such as stripping or merging. In this circumstance, we reassign a brand new QSO elsewhere that is sampled from the current redshift's QXLF.

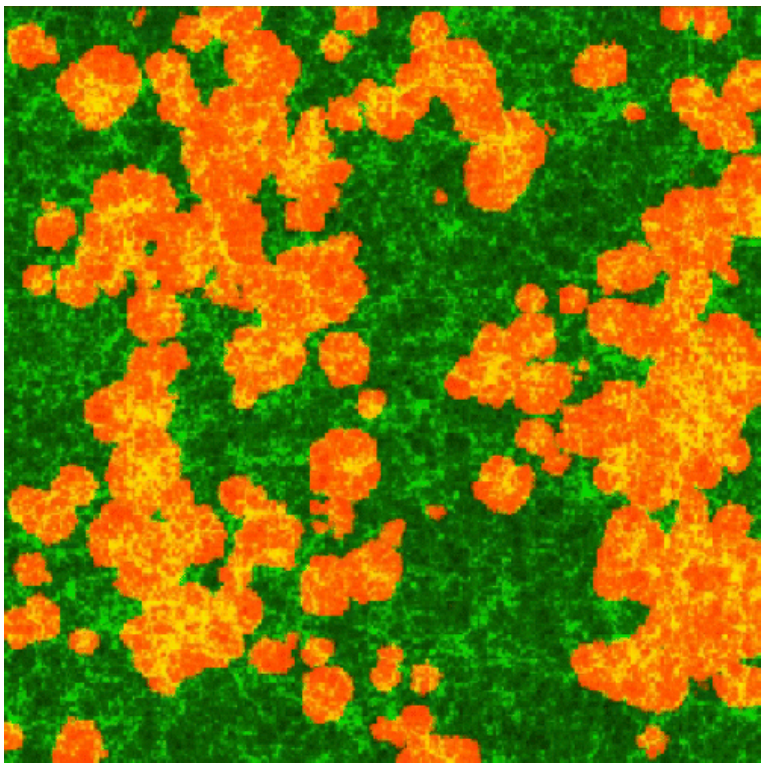

Figure 2.4: Plot from Mellema et al. (2006b) of the ionized bubbles (orange) and neutral regions (green) obtained from C^2-Ray. The underlying density field was produced by cubep^{3}m. The box is 143 $\times$ 143 Mpc.

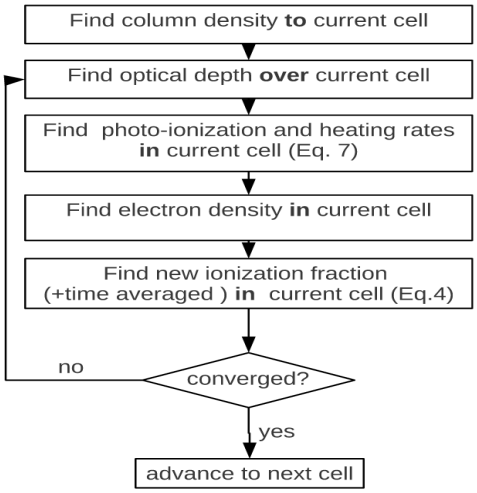

Figure 2.5: Flow chart showing the algorithm employed by the original version of C^2-Ray (Mellema et al., 2006a) for modelling the process of reionization by UV photons.

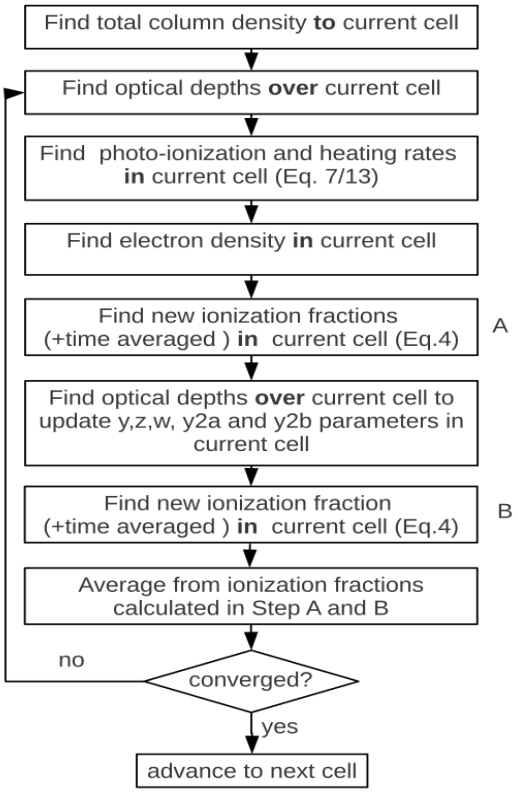

Figure 2.6: Flowchart of the updated method from Friedrich et al. (2012). This version of the code incorporates the effects of hard radiation and tracks temperature.

1) The evolution of the ionization of a cell is assumed to follow an exponential decay converging on the equilibrium solution. When a cell is optically thick the photo-ionization rate is set to be consistent with the number of photons absorbed within that cell. This method is able to correctly track the progress of ionization fronts passing over many cells within one single timestep and has been presented in Mellema et al. (2006a). This allows cells to be far larger than in other approaches.

2) The value of the optical depth is taken to be the time-averaged value throughout the timestep. When this value is used in conjunction with the finite-volume approach the timesteps can be increased - making C^2-Ray orders of magnitude faster than more standard numerical approaches.

The outline of the algorithm is displayed in Figure 2.5. Firstly, the column density is

calculated and used to find the time averaged optical depth for all species, $< \tau_{\nu,\mathrm{i}} >$. The photoionization rate of each cell, Γ_j, is calculated as follows:

$$\Gamma = \int_{vth}^{\infty} \frac{L_\nu e^{-<\tau_\nu>}}{h\nu} \frac{1 - e^{-<\tau_\nu>}}{< n_{HI} > V}, \tag{2.6}$$

where L_ν is the source luminosity, h is the Planck constant, $< n_i >$ is the number density of hydrogen, V is the volume of a hollow sphere where the inner and outer radii corresponds to the distance between the source and closer and further side of the cell respectively and $< \tau_\nu >$ is the time averaged optical depth. $< \tau_\nu >$ is the sum of all the time averaged optical depths of the cells between the current cell and the source: $< \tau_\nu >= \sum_c < \Delta\tau_{\nu,\mathrm{c}} >$ where:

$$< \Delta\tau_{\nu,\mathrm{c}} >= n_{\mathrm{HI}} \sigma_\nu \Delta r (1 - < x >), \tag{2.7}$$

where σ_ν is the cross section for the ionizing photons, Δr is the width of the cell and $< x >$ is the time averaged ionized fraction. σ_ν is taken to be a constant as Δr is very small and gray opacity can be assumed. $< x >$ is calculated using:

$$< x >= x_{\mathrm{eq}} + (x_0 - x_{\mathrm{eq}})(1 - e^{-\Delta t/ti}) \frac{ti}{\Delta t}, \tag{2.8}$$

where x_0 is the initial value of x, Δt is the timestep and x_{eq} is given by:

$$x_{\mathrm{eq}} = \frac{\Gamma + n_e C_H}{\Gamma + n_e C_H + n_e \alpha_H}, \tag{2.9}$$

and t_i is given by:

$$t_i = \frac{1}{\Gamma + n_e C_H + n_e \alpha_H}. \tag{2.10}$$

Here n_e is the number of electrons, C_H is the collisional ionization rate, and α_H is recombination coefficients for hydrogen. Both C_H and α_H are temperature dependant.

In order to solve Equation 2.6 an iterative approach must be employed as both sides of the equation depend on Γ. This is where high performance computation becomes necessary as many iterations are required. In order to accelerate the process Γ can evaluated at the hydrogen ionization threshold ($\tau_0 = \tau_{\nu=\nu_t hr}$):

$$\Gamma = \Gamma'(\tau_0) + \Gamma'(\tau_0 + \Delta\tau) \tag{2.11}$$

, where

$$\Gamma'(\tau_0) = \int_{vth}^{\infty} \frac{L_\nu e^{-<\tau_\nu>}}{h\nu < n_{\mathrm{HII}} V}. \tag{2.12}$$

To avoid expensive integrals at runtime values from Equation 2.12 are pre-calculated and stored in photo-ionization tables and found via interpolation at runtime.

Note that when helium is introduced the photo-ionization rate for all species is calculated as the sum over all frequency bins j:

$$\Gamma = \sum_j \frac{\tau_j}{\tau_j^{\text{total}}} \frac{\Gamma_j}{n} \qquad i = \text{HI, HeI, HeII,} \tag{2.13}$$

where n_i is the number density of species i.

The recombinations of hydrogen and (helium I in the Helium version) are calculated using a fit obtained by Aldrovandi and Pequignot (1973), assuming the temperature of the ionized region to be 1×10^4K. Recombinations to helium 2 are calculated using a fitting formula from Osterbrock and Ferland (2006), again assuming a temperature of 1×10^4K. Photons released by these recombinations can be energetic enough to ionized other atoms. Such photons are treated using the on the spot approximation, i.e. they are considered to be absorbed locally.

Next the evolution of the ionized fraction of hydrogen is obtained from the chemical evolution equations:

$$\frac{d}{dt}\begin{pmatrix} x_{\text{HI}} \\ x_{\text{HII}} \end{pmatrix} = \begin{pmatrix} -\Gamma_{\text{HI}} + C_{\text{HI}} n_e & \alpha^{\text{B}}_{\text{HII}} n_e \\ \Gamma_{\text{HI}} + C_{\text{HI}} n_e & -\alpha^{\text{B}}_{\text{HII}} n_e \end{pmatrix} \begin{pmatrix} x_{\text{HI}} \\ x_{\text{HII}} \end{pmatrix}, \tag{2.14}$$

where x_{HI} is the neutral hydrogen fraction, x_{HII} is the ionized hydrogen fraction, n_e is the electron density, Γ_{HI} is the hydrogen photo-ionization rate, C_{HI} is the collisional ionization rate and $\alpha^{\text{B}}_{\text{HII}}$ is the recombination rate.

2.3.1 Heating Simulations

C^2-Ray was improved in order to accommodate harder sources by including multi-frequency RT in (Friedrich et al., 2012). The effects of helium, secondary ionizations, multi-frequency photo-ionization and detailed heating through the full on-the-spot approximation were added to the basic algorithm. Three frequency bands are used for the photo-ionization and photo-heating rates: a band encompassing the frequencies able to ionize hydrogen only, a band for photons that can ionize both hydrogen and helium once and finally a band that is capable of all possible ionizations. These bands are split into 1, 26 and 20 sub-bins respectively and the convergence for these numbers has been tested in Friedrich et al. (2012). These rates are pre-calculated and stored in lookup tables that are functions of ionized fraction and optical depth.

The method of calculating the photo-ionization rates is slightly altered. Now, the optical depth depends on three species: HI, HeI and HeII. Therefore we must sum over these

species:

$$\tau^{\text{total}}(\nu) = \sum_j \tau^{\text{total}} \left(\frac{\nu}{\nu_j}\right)^{-\eta_j}. \tag{2.15}$$

Note that this assumes that the optical depth of all three species follow the same power law index η_j within the frequency bin. The rate at which ionizing photons are passing through the cell, $\mathcal{G}_j^{\text{total}}$, is then found as follows:

$$\mathcal{G}_j^{\text{total}} = \frac{G_j\left(\tau_{\text{in}}^{\text{total}}\right) - G_j\left(\tau_{\text{out}}^{\text{total}}\right)}{V}, \tag{2.16}$$

The calculation then proceeds as before until we reach the photo-ionization rate, which is now calculated for species i as the sum over all frequency bins j:

$$\Gamma^i = \sum_j \frac{\tau_j^i}{\tau_j^{\text{total}}} \frac{\mathcal{G}_j^{\text{total}}}{n_i}, \tag{2.17}$$

where n_i is the number density of species i.

To calculate the temperature change we must calculate the energy change per molecule $\delta E/\delta t = \mathcal{H} - C$, where $\mathcal{H}$ is the heating rate and C is the cooling rate. Photo-heating (where additional energy left over from photo-ionizations is absorbed into the gas) is the only heating mechanism considered in this work. Hence, the heating rate is given by:

$$\mathcal{H} = \sum_i \frac{\tau_{\text{i}}}{\tau_{\text{tot}}} \mathcal{H}(i), \tag{2.18}$$

for species i where τ is the optical depth, $\mathcal{H}(i)$ is given by:

$$\mathcal{H}(i) = \int_{\nu_{\text{sub-bin}}} \frac{h(\nu - \nu_{\text{th}}(i)) L_\nu e^{-<\tau_\nu>}}{h\nu} \frac{1 - e^{-<\Delta\tau>}}{\nu_{\text{shell}}} d\nu \tag{2.19}$$

where ν is the frequency, τ is the optical depth and L_ν is the Luminosity as frequency ν.

The adiabatic expansion of the Universe and Compton cooling from cosmic microwave background photons are included along with cooling from the free-free interactions, recombinations and collisional excitation of all specials. For more details on these cooling processes see Friedrich et al. (2012).

After an atom has been ionized, the free electron can retain enough energy to ionize another atom in a future collision, known as a secondary ionization. The energy used in this process does not contribute to heating the gas. It is therefore important to include this process in order to obtain the correct temperature. It is not computationally feasible to fully include this process in these cosmological simulations. Therefore, Friedrich et al. (2012) implemented the separable functional relationship found by Ricotti et al. (2002).

During this work we found that the recombination rates were incorrectly calculated. This was due to the fact that they were calculated from the average temperature of the cell, which in cosmological simulations is often partially ionized, due to the kilo parsec scales of ionization fronts being unresolvable with current day computational power. These partially ionized cells not only led to the incorrect temperature and ionized fraction, but caused $\delta T_{\rm b}$ to be vastly overestimated. This over estimation arises from the fact that the 21-cm signal is only emitted from neutral gas. However, the version of C^2-Ray presented in Friedrich et al. (2012) calculates the average temperature of the entire region, including the hot ionized gas (which does not contribute to the signal).

As a part of this work, a post-processing method was developed to extract the temperature of the neutral gas (described in Appendix A) and applied to the simulations presented in Chapter 3. These corrections were found to be insufficient for cases of late X-ray heating, such as those presented in Chapter 4, and the code itself was updated in order to obtain these results (see Appendix C). All simulations presented in this work are correct to the best of our knowledge.

Finally, the code was further updated to include QSO sources as well has HMXB sources. Unlike HMXBs the QSOs can turn on and off, and the luminosity is given as an input in the source file rather than being decide by the underlying mass. See Appendix B for more details on this.

The addition of heating can impose additional constraints on the time-stepping, meaning smaller time-steps are required Lee et al. (2016). However, for the spacial and temporal resolutions used in our suite of simulations are converged.

2.4 Radiative Transfer of Soft UV Photons

2.4.1 Lyman-α

Fully numerical radiative transfer simulations of Lyman α are computationally expensive and, in these big simulation volumes, largely unnecessary. The profile $1/\mathrm{R}^2$ (where R is radial distance from the source) has been shown to be consistent with detailed radiative transfer simulations in Semelin et al. (2007); Vonlanthen et al. (2011); Higgins and Meiksin (2012) on large scales. In addition Lyman α photons from a point source are thought to produce nearly spherical profile, even in the presence of density fluctuations (Vonlanthen

et al., 2011).

To calculate the Lyman-α coupling coefficients, y_α we use an improved version of the method presented in Ghara et al. (2016) assuming the $1/R^2$ relation. The method was updated to be able to take into account the changing halo mass, and therefore luminosity, of the sources. In addition, it was improved to be able to model the Lyman-α from sources that move around or are suppressed. In order to calculate the Lyman-α flux, J_α from the simulations, we use the spectra outlined in section 2.2.1.

Chapter 3

Simulating the Impact of X-ray Heating during the Cosmic Dawn

3.1 Contributions of the Authors

The work presented in this chapter has been previously published in Ross et al. (2017). The contributions of the author of this paper to the work:

- **Hannah Ross:** ran the majority of the simulations, create and apply a post-processing method to retreive the 21-cm signal from the simulations in this chapter (see Appendix A), to calculate the statistics, produce all plots for the paper and paper writing.
- **Keri Dixon:** Contributed significantly to the development of the post-processing method and paper writing.
- **Ilian Iliev:** Ran one of the simulations (however, it a bug in the Helium rates was found and it needed to be re-run), helped develop the post-processing method, paper writing, secured computing time and provided general guidance.
- **Garrelt Mellema:** helped develop the post-processing method and contributed to paper writing.

3.2 Version of C²-Ray used in this Work

During this work, the inadequacies of the heating algorithm in C²-Ray described in section 2.3.1 became apparent and a method to account for this was developed. The results of the simulations presented in this chapter were all post-processed in order to obtain the correct temperature of the neutral gas.

These calculations are described in detail in Appendix A along with the results from test boxes showing the method's effectiveness. As described in section 2.3.1 the recombination rates are calculated incorrectly. However, due to the small size of the ionized regions at this time this does not impact the 21-cm signal until significant reionization occurs (z≈10).

3.3 Introduction

In this Chapter, we focus on work done on the fluctuations in δT_{rmb} due to inhomogeneous heating from HMXBs. We present the first fully numerical large scale (349 comoving Mpc per side) simulations ever run of inhomogeneous X-ray heating during the Cosmic Dawn which include the effects of helium and multi-frequency heating. For reference we also include results from a simulation run without X-ray sources.

The density fields and source catalogues are obtained from CUBEP³M (see section 2.1 for further details). The sources used in this work have been described in detail in section 2.2 and the radiative transfer is rung run using the code C²-Ray (see section 2.3). We assume that the Lyman-α background has been built up sufficiently that the signal is completely decoupled from the CMB and that collisional coupling is insignificant.

Two simulations are run, one with HMXBs and one without. Both simulations contain stellar sources, that contribute to ionization but do not produce photons energetic enough to penetrate the IGM (see section 2.2.1). HMXBs are responsible for the heating of the IGM in the X-ray simulation (see section 2.2.2). In this work we focus on the impact of the inhomogeneous X-ray heating on the 21-cm signal. In particular we investigate the fluctuations produced by our X-ray model and their detectability by SKA1-Low.

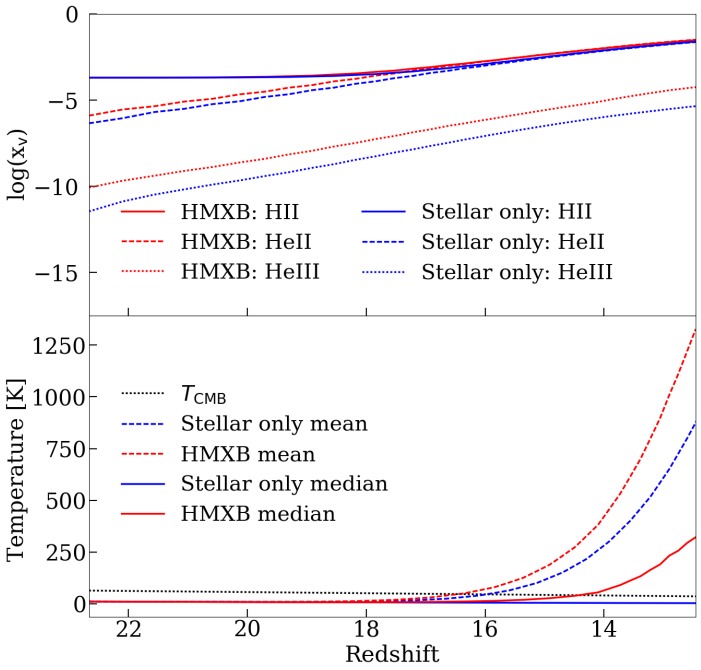

Figure 3.1: (top) The mean ionized fraction by volume of each species: HII - solid lines, HeII - dashed lines and HeIII - dotted lines. (bottom) The volume-weighted mean temperature - dashed lines, median temperature - solid lines and T_{CMB} - dotted line. In both plots the HMXB case is shown in red and the stellar-only case is shown in blue.

3.4 Reionization and Thermal Histories

The 21-cm signal is affected by the thermal and ionization histories of the IGM. In the upper panel of Fig. 3.1, we show the mean volume-weighted ionized fraction evolution for the species present in our simulations: H II (solid lines), He II (dashed lines), and He III (dotted lines) for the HMXB case (red lines) and the stellar only case (blue lines). The ionization of H I and He I is largely driven by the hard-UV photons of stars, which are more abundant than X-ray photons. The effect of the latter is largely limited to increasing the He III abundance by about an order of magnitude (while remaining low) due to their high energy per photon.

In the lower panel, volume-weighted mean (dashed lines) and the median (solid lines) temperatures are shown for the HMXB case (red lines) and the stellar-only case (blue lines). For the HMXB case, the mean temperatures are increased modestly, surpassing $T_{\rm CMB}$ earlier than in the stellar-only case, with both occurring around $z \sim 16$. The

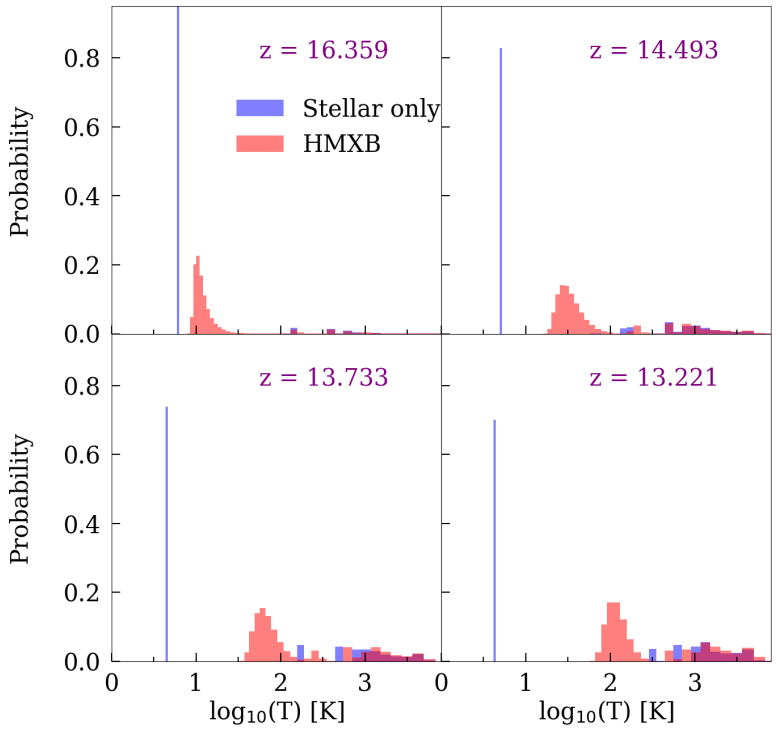

Figure 3.2: Histograms of the temperature at the full simulation resolution for the HMXB (red) and stellar-only (blue) cases for several illustrative redshifts.

differences between the mean $T_{\rm K}$ of the two cases (with and without X-rays) grow at later times, rising above 50 per cent for $z < 13$. As the 21-cm signal probes neutral hydrogen, the median temperature of the IGM is more relevant, since the mean is skewed towards higher value by the hot, ionized regions. In the presence of X-rays, the median surpasses $T_{\rm CMB}$ just before $z = 14$; while in their absence the neutral IGM remains cold. These temperatures are somewhat higher than those found in Madau and Fragos (2017), which is likely due to the fact our spectra is a single power-law spectra whereas their SED more similar to a double power law and as a result has less low energy photons. The long mean free paths of higher energy photons means that they heat gas much later than they are emitted or not at all.

The only previous full numerical simulations to take into account X-ray sources are from Baek et al. (2010)[1]. However, the mass resolution of these simulations is approximately 600 times lower than ours, and the volume is 15 times smaller. As a consequence, their

[1]The simulations in Xu et al. (2014) and Ahn et al. (2015) are focused on a single X-ray source in a zoomed region of a cosmological volume, so are not directly comparable to our results here.

first sources of any type of radiation appear around $z \approx 14$, approximately when the neutral IGM in our simulation has already been globally heated to well above the $T_{\rm CMB}$. In other words, their simulations describe a scenario in which X-ray sources appear very late, are relatively rare and bright, and are coincident with substantial stellar emission. This situation is very different from our case in which the first X-ray sources appear at $z \approx 23$ and large numbers of relatively faint sources heat the neutral IGM well before any substantial reionization. We will, therefore, not further compare the details of our results to those of Baek et al. (2010).

More detailed information about the temperature distributions is obtained from the corresponding probability distributions functions (PDFs), shown in Fig. 3.2. Again, the HMXB case is shown in red and the stellar-only case in blue. These PDFs were generated from the coeval simulation cubes using 100 bins and normalised to have a total area of one. The stellar-only distributions are clearly bimodal, with a few hot, partially ionized regions and the majority of cells remaining very cold. This behaviour is expected, given the very short mean free path of the ionizing photons in this case, which yields sharp ionization fronts. In contrast, when X-rays are present, their long mean free paths lead to gas heating spreading quickly and widely, with all cells being affected. The distribution is strongly peaked, relatively wide, and gradually moves towards higher temperatures, with typical values above 100 K by $z = 13.2$. Our thermal history is similar to that of Case A (Pop. II stars) in Pritchard and Furlanetto (2007) and case '$\log \zeta_{\rm X} = 55$' in Watkinson and Pritchard (2015). These studies do not provide temperature PDFs or median values.

When compared to the emissivities in the 2-10keV range, ϵ_{2-10}, of the HMXBs used in Madau and Fragos (2017) our values are an much higher. They predict a value of $\epsilon_{2-10} \approx 10^{38}$ ergs s^{-1} comoving Mpc^{-3} for z$\sim$10. The ϵ_{2-10} for our HMXBs start off much lower, at $\epsilon_{2-10} \approx 10^{36}$ ergs s^{-1} comoving Mpc^{-3} and rise to $\epsilon_{2-10} \approx 10^{39}$ ergs s^{-1} comoving Mpc^{-3} by z$\sim$13. It should be noted that our HMXBs only extend to 5.4 keV so would be much brighter than this in the 2-5keV range.

The lightcones (spatial-redshift/frequency slices) provide a visual representation of the quantities discussed above, including spatial variations and evolution over time (Fig. 3.3). These lightcones are constructed by taking a cross section of the simulation volume along the line of sight and continually interpolating in time the relevant quantity using the spatial periodicity of our cosmological volume.

In Fig. 3.3 (top panels), we show the hydrogen ionization lightcone. As expected based

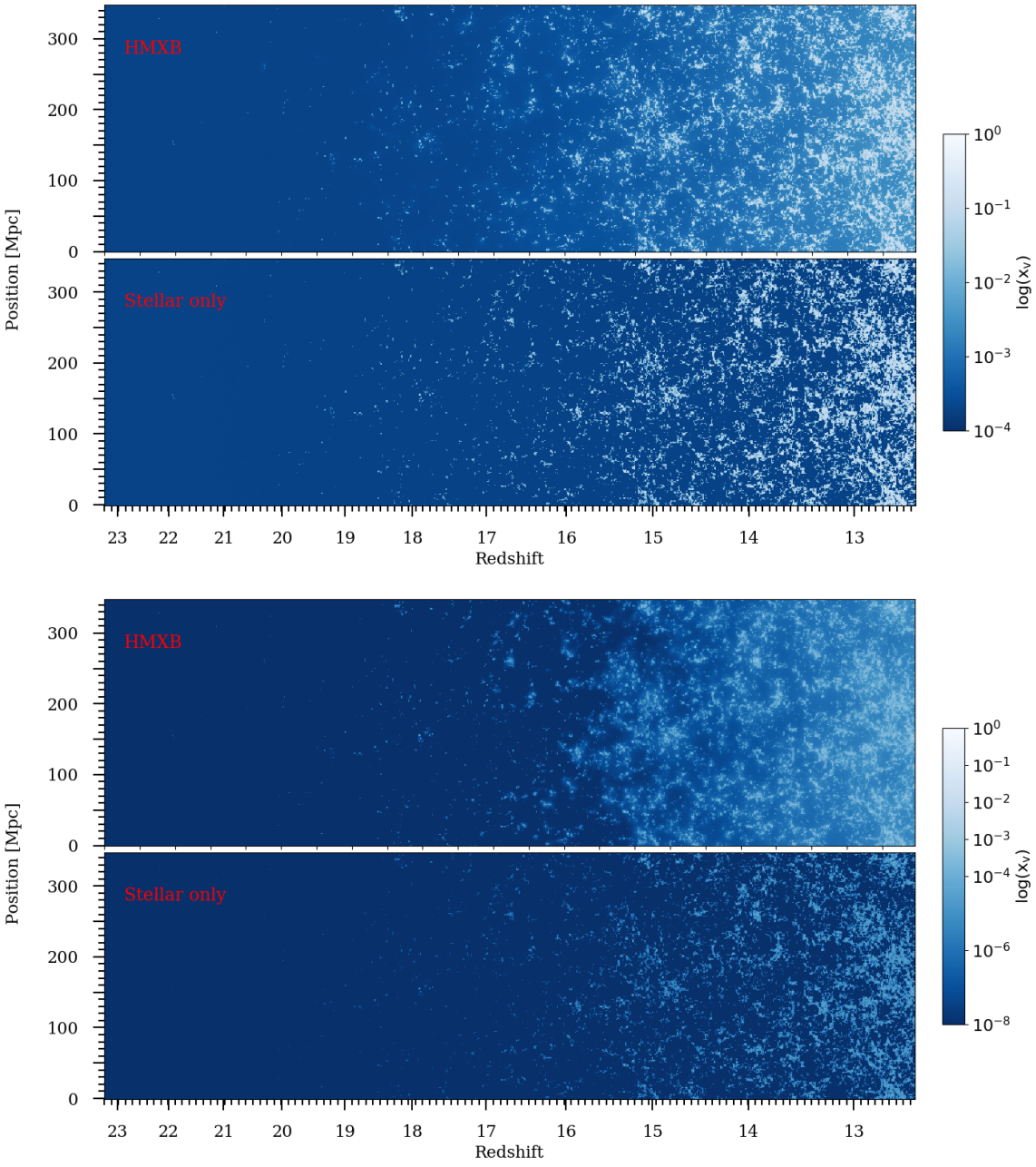

Figure 3.3: The position-redshift lightcone images of the ionized volume fraction of hydrogen (top two panels) and He III (bottom two panels). Shown are the HMXB case (top panel in each pair), and the stellar-only case (lower panel in each pair).

on the very similar mean ionization fractions in the two simulations, the morphology of hydrogen ionization is broadly similar. However, the hard photons, which penetrate deep into the neutral regions produce a low-level, but widespread ionization of the IGM and 'fuzzier', less clearly defined H II regions when X-ray sources are present. Given their similar ionization potentials, the first ionization of helium (not shown here) closely follows that of hydrogen. Alternatively, the second helium ionization potential is sufficiently higher to result in significant differences between models (Fig. 3.3, bottom panels). The 50,000 K blackbody stellar spectra produce very few photons able to fully ionize a helium atom; thus, any He III produced is concentrated in the immediate surroundings of the ionizing sources. However, the X-rays are very efficient in fully ionizing helium, producing widespread ionization (albeit still at a relatively low level). This ionization is also quite patchy on large scales, especially at early times. The exact morphology depends on the spectra, abundance, and clustering of the X-ray sources.

The lightcones in Fig. 3.4 (top panels) show the spatial geometry and evolution of the IGM heating. The soft, stellar radiation in both models ionize and heat the immediate environments of the sources to $T \sim 10^4$ K, seen as dark regions, with the majority of the IGM remaining completely cold. The X-ray radiation propagates much further, starting to heat the gas throughout. Large, considerably hotter regions (black in the image) develop, quickly reaching tens of Mpc across before $z \sim 15$. These regions gradually expand and merge, resulting in thorough heating to hundreds of degrees by $z \sim 14$; though, large cold regions still remain present down to $z \sim 13.5$. At early times, the heating from X-rays increases the inhomogeneity when X-rays are present, but the temperature distribution eventually becomes more homogeneous at later times.

3.5 21-cm Differential Brightness Temperature

We are primarily interested in the directly observable quantity of this epoch, the 21-cm $\delta T_{\rm b}$. As discussed in section 1.3, $\delta T_{\rm b}$ depends on the density, ionization, and temperature fields. In Fig. 3.4 (bottom panels), we show the $\delta T_{\rm b}$ lightcones corresponding to the same cross section through the position-redshift image cube as in Fig. 3.3 and Fig. 3.4 (top panels). The morphology of the $\delta T_{\rm b}$ fluctuations is closely related to that of the heating, demonstrating the importance of temperature variations, especially during the early stages of the EoR. Long-range X-ray heating produces a gradual, extended transition from absorption into emission. Large-scale fluctuations are significant throughout, and the

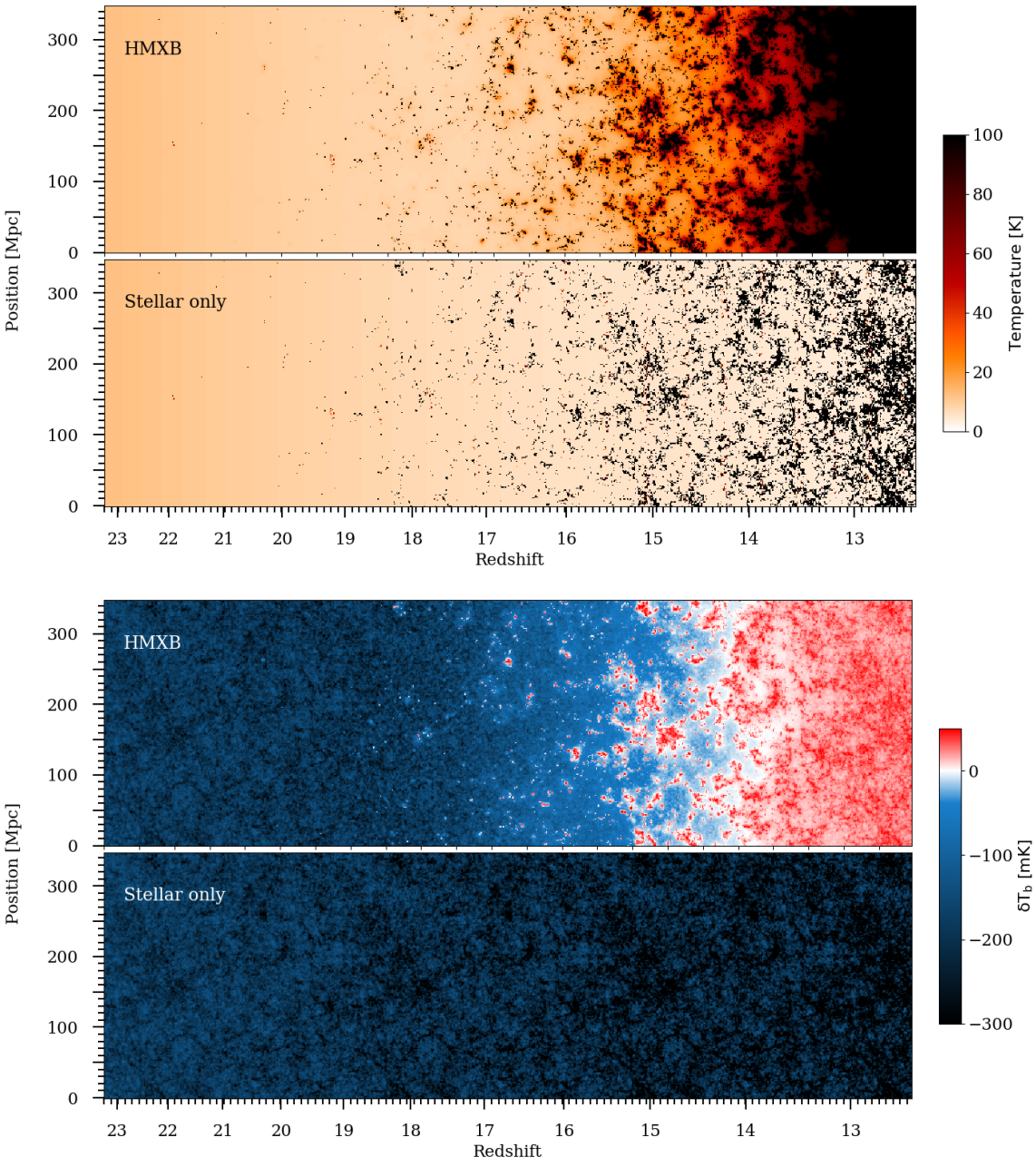

Figure 3.4: The position-redshift lightcone images of the IGM gas temperature (top two panels) and 21-cm differential brightness temperature (bottom two panels). Shown are the HMXB case (top panel in each pair), and the stellar-only case (lower panel in each pair).

first emission regions appear part way through the simulation at $z = 18$, after enough X-rays have penetrated into the IGM near the ionizing sources in order to heat it. After this these bubbles grow quickly, with some reaching tens of Mpc in size by $z \sim 15.5$. Only after $z \sim 13.8$ have all large regions of 21-cm absorption disappeared.

Comparing to Mesinger et al. (2013), the results of our simulation appear to be closest to their case with X-ray efficiency $f_{\rm x} = 1$. They also find an extended transition from absorption to emission starting at $z \sim 20$, which completes around $z \sim 14$. However, by that time, the hydrogen ionization fraction is around 10 per cent, which is substantially higher than our case. The majority of the difference is likely due to the fact that their sources are more efficient, as indicated by a completion of reionization by $z_{\rm reion} \sim 8$ in their case versus $z_{\rm reion} < 6.5$ for our source model (Dixon et al., 2016b).

In the stellar-only case, the signal remains in absorption throughout the simulation as there are no photons with long enough mean free paths to penetrate and heat the neutral IGM. As a result the IGM simply cools adiabatically as the universe expands[2]. In reality the temperature of the IGM would also be impacted by shock heating, however, this has not been included in our simulations and is left for future work.

In Fig. 3.5, we show the power spectra of $\delta T_{\rm b}$ at several key redshifts. During the early evolution (shown is $z = 20.134$), only a modest amount of heating of the IGM has yet occurred in the X-ray model. Thus, the large-scale 21-cm fluctuations are dominated by the density variations, which are the same in the two cases. Therefore, the 21-cm power spectra are almost identical at this stage, with power suppressed slightly on all scales in the HMXB case. The high-$T_{\rm K}$ limit results in significantly lower fluctuations, reflecting the lower average $\delta T_{\rm b}$ in the emission regime compared to the absorption or mixed regimes.

As the evolution progresses and with X-ray sources contributing to the long-range heating, the 21-cm power is significantly boosted on large scales and a well-defined, if broad, peak develops around scale of $\sim$43 Mpc ($z = 15 - 16$). The small-scale power decreases due to the stronger heating in the vicinity of sources, which brings the temperature contrast with the CMB down and closer to the high-$T_{\rm K}$ limit. The overall fluctuations peak at ~ 14 mK around $z \sim 15$ in the HMXB model. Pritchard and Furlanetto (2007) find a heating peak that is in agreement with our result, at similar scales ($k \sim 0.14\,{\rm Mpc}^{-1}$) with an amplitude

[2]There is also some Compton heating due to CMB scattering, which we do include in our simulation. This heating is inefficient at this point due to the low density of the IGM.

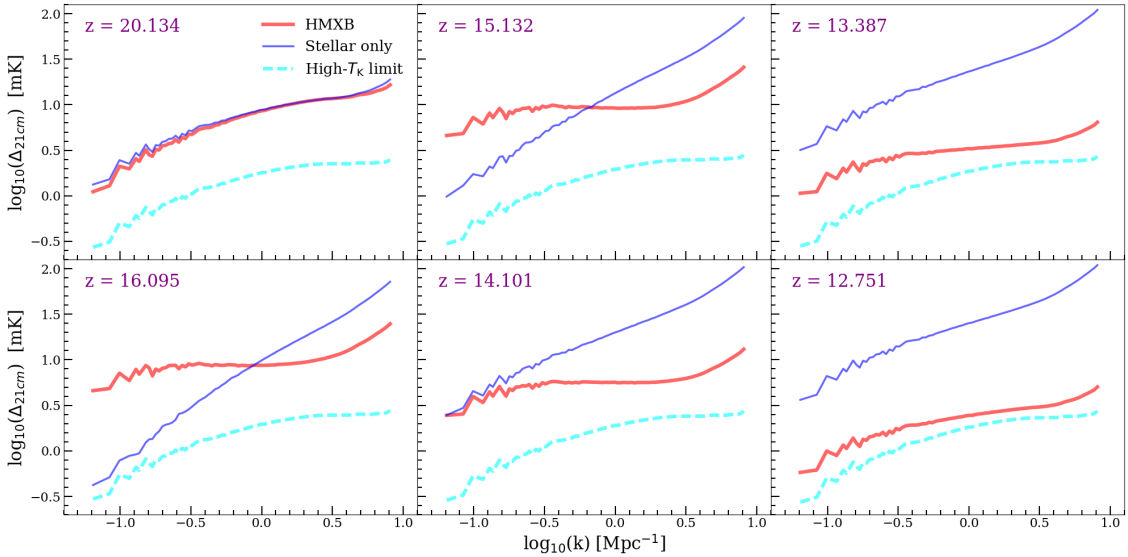

Figure 3.5: The 21-cm power spectra from our simulations at several key stages of the evolution with the high-$T_{\rm K}$ limit results for reference. The high-$T_{\rm K}$ limit is shown in yellow and as before the results from the HMXB case are shown in red and the stellar-only case in blue.

of$\Delta_{21\rm cm} \sim 11.5$mK or $\Delta_{21\rm cm} \sim 20$ mK depending on the source model used. Results from Pacucci et al. (2014a) are also in rough agreement with our own, with a power spectra peak at a similar scale (44 Mpc) and at similar redshift ($z \sim 15 - 16$). Their peak value is also in agreement at $\Delta_{21\rm cm} \sim 14\,$mK.

Over time, as emission patches develop and the overall IGM is gradually heated ($z = 12 - 14$), the power spectra slowly approach the high-$T_{\rm K}$ limit, but does not reach it even by the end of our simulation. At that point ($z = 12.7$), the IGM has been heated well above $T_{\rm CMB}$ throughout and therefore is in 21-cm emission everywhere. However, the neutral IGM temperature remain at only a few hundred degrees, is still considerably spatially inhomogeneous, and so is not yet fully in the high-$T_{\rm K}$ limit, where the $\delta T_{\rm b}$ becomes independent of the actual gas temperature value.

The evolution is markedly different in the stellar-only case. The 21-cm fluctuations here are dominated by density fluctuations as the stars do not produce many photons able to penetrate into and heat the neutral cold IGM. Consequently, the shape of the power spectrum remains almost a power law. Throughout the simulation cosmic structures continue to form so the amplitude of these fluctuations gradually increases over time. In the stellar-only case, 21-cm power on all scales is far higher at late times compared to that in the HMXB case since the mean temperature in the latter case approaches and then

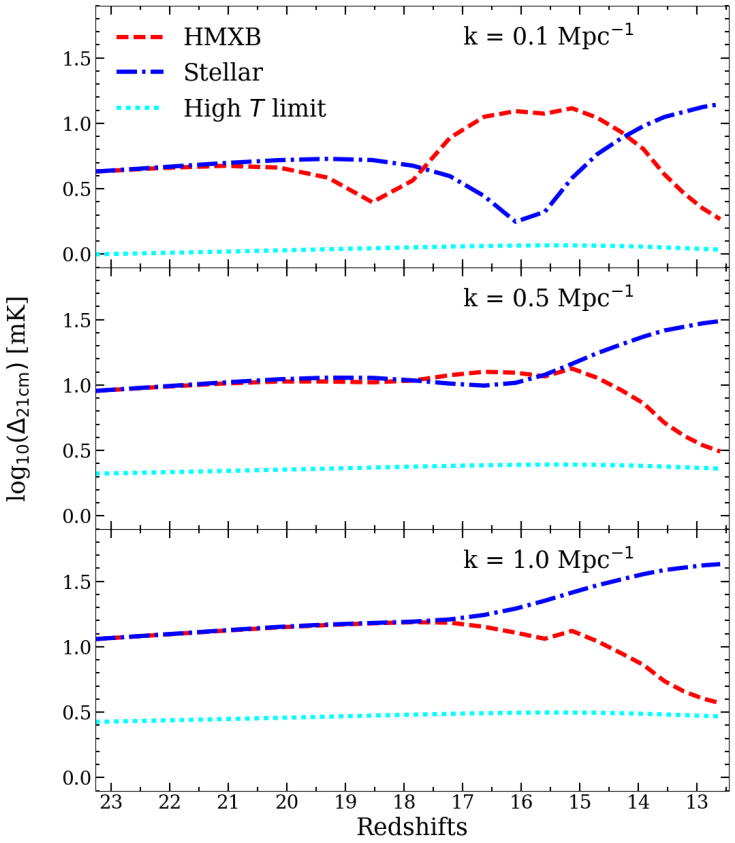

Figure 3.6: The evolution of the 21-cm power spectra modes at $k = 0.1, 0.5$ and $1\,\mathrm{Mpc}^{-1}$ for the two simulations and high-T_K limit for reference, as labelled.

surpasses $T_{\rm CMB}$ and the IGM remains very cold in the former.

The evolution of several particular k-modes ($k = 0.1, 0.5$, and $1\,{\rm Mpc}^{-1}$) is shown in Fig. 3.6. With increasing k, the 21-cm fluctuations deviate from the density fluctuations later due to more advanced structure formation at small scales. With X-rays present, there is a peak from heating at all scales considered here, occurring between $z \sim 15 - 17$. The peak becomes wider and more pronounced at larger scales, which matches the typical 21-cm fluctuations scale due to inhomogeneous heating (on the order of tens of Mpc). At larger scales, X-ray heating from the first sources removes the colder regions close to these sources, resulting in an initial dip in the power. Subsequently, the power rises as the X-ray heating extends inhomogeneously into the IGM. Both of these features are present, but less pronounced, at the intermediate scale (k=$0.1\,{\rm Mpc}^{-1}$). In the absence of X-rays, the evolution is markedly different. At later times when the IGM X-ray heating has become homogeneous, the corresponding result for the stellar-only case has higher power on all scales due to the IGM remaining cold. In the high-$T_{\rm K}$ limit, the fluctuations are much lower and mostly flat at all scales due to the lack of cold IGM, which results in lower amplitude signal driven initially by the density fluctuations only.

The evolution of the $k = 0.1\,{\rm Mpc}^{-1}$ mode is roughly in agreement with Mesinger et al. (2013), however, their peak occurs earlier than our results, again likely due to the higher assumed efficiency of their sources. Probably for the same reason, Pritchard and Furlanetto (2007) and Fialkov et al. (2014) find the peak to occur later. Beyond the timing of the peak, the amplitude is in rough agreement with in the semi-numerical results. Mesinger et al. (2013), Pritchard and Furlanetto (2007), and Santos et al. (2010) find a marginally higher peak value of 20 mK, while Fialkov et al. (2014) find a lower one of 7 mK.

In Fig. 3.7, we show maps of the mean-subtracted $\delta T_{\rm b}$ ($\delta T_{\rm b} - \overline{\delta T_{\rm b}}$) at several key epochs of the redshifted 21-cm evolution, smoothed with a Gaussian beam that is roughly twice as broad as the what can be achieved by the core of SKA1-Low, which will have maximum baselines of around 2 km. We averaged over a frequency bandwidth that corresponds to the same spatial scale as the Gaussian beam. To mimic the lack of sensitivity at large scales, which interferometers have due to the existence of a minimum distance between their elements, we also subtracted the mean value from the images. We do not include instrument noise and calibration effects in these maps. However, at this resolution, SKA1-Low is expected to have a noise level of ~ 10 mK per resolution element for 1000 hours of integration (Koopmans et al., 2015).

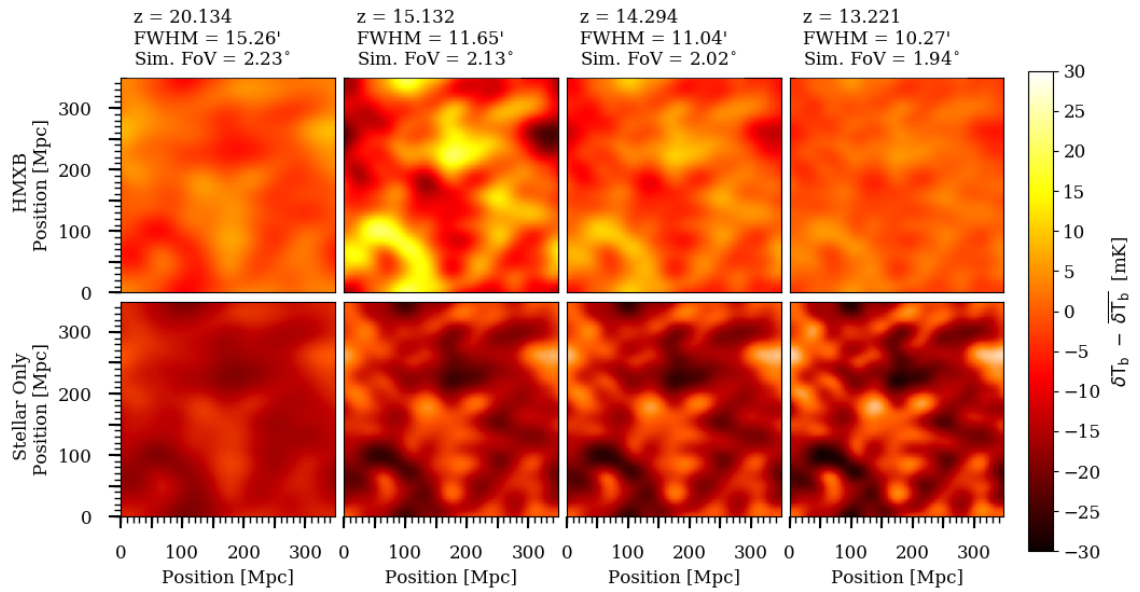

Figure 3.7: Mean-subtracted δT_b maps smoothed with a Gaussian beam with the FWHM corresponding to a 1.2 km maximum baseline at the relevant frequency, as labelled. The images are bandwidth-smoothed with a top hat function (width equal to the distance corresponding to the beam width). The X-ray simulation runs along the top row and the stellar-only case is below, with snapshots of the same redshifts being vertically aligned.

The images show a clear difference between the cases with and without X-ray heating even at these low resolutions, suggesting that SKA should be able to distinguish between these two scenarios. As the variations over the field of view (FoV) reach values of 50 mK we can also conclude that SKA1-Low will be able to image these structures for deep integrations of around 1000 hours. Previous expectations were that SKA1-Low would only be able to make statistical detections of the 21-cm signal from the Cosmic Dawn. Our results indicate that, at least from the perspective of signal to noise, imaging should be possible.

Once again, the signal from partially heated IGM peaks around $z \sim 15$, when there are large regions – tens of Mpc across – in either emission or absorption. This peak is followed by gradual thorough heating of the IGM above T_CMB, bringing the signal into emission and, thus, decreasing the overall fluctuations. In the stellar-only case, the maps remain fully in absorption at these resolutions, since the ionized regions are much smaller than the beam extent and are smoothed away. Nonetheless, considerable fluctuations remain, with high overall amplitude due to the very cold IGM in that case. Over time larger regions are substantially affected by further structure formation occurs and the contrast in the images becomes greater.

3.6 21-cm One-Point Statistics

The 21-cm fluctuations are typically non-Gaussian in nature and are not fully described by the power spectra alone, and imaging might not be possible in all regimes for sensitivity reasons. Therefore, the one-point statistical properties of the 21-cm signal are also of great interest, since they quantify other aspects of the 21-cm signal and enable comparisons with past works and future observations.

These quantities are calculated from coeval simulation cubes, smoothed with a Gaussian beam corresponding to a 1.2 km maximum baseline at the relevant frequency and a bandwidth corresponding to the same spatial extent as the full-width half-maximum (FWHM) of the beam. This smoothing is the same as was used for the images in Fig. 3.7.

The rms is defined as:

$$\mathrm{rms}(y) \equiv \sigma = \sqrt{\frac{\sum_{i=0}^{N}(y_i - \overline{y})^2}{N}}, \tag{3.1}$$

and we use the dimensionless definitions for the skewness and kurtosis, as follows:

$$\mathrm{Skewness}(y) = \frac{1}{N}\frac{\sum_{i=0}^{N}(y_i - \overline{y})^3}{\sigma^3}, \tag{3.2}$$

and

$$\mathrm{Kurtosis}(y) = \frac{1}{N}\frac{\sum_{i=0}^{N}(y_i - \overline{y})^4}{\sigma^4}. \tag{3.3}$$

Here, y is the quantity of interest (in this case δT_b), N is the number of data points (i.e. the total number of cells in smoothed simulation results), $\overline{y}$ is the mean value of y, and σ^2 is the variance of y.

In the top-left panel of Fig. 3.8, we show the global value of the δT_b calculated as the mean signal from our simulations, $\overline{\delta T_\mathrm{b}}$. The high-$T_\mathrm{K}$ limit corresponds to the dashed (cyan) lines. In both simulation models, the global value starts negative, due to the initially cold IGM, and drops further as the universe expands and cools adiabatically. In the stellar-only case (thin, blue line), $\overline{\delta T_\mathrm{b}}$ rises slowly thereafter, starting at $z \sim 16.5$ as the highest density peaks become ionized. The highest value remains negative, since the neutral IGM never gets heated in this scenario. $\overline{\delta T_\mathrm{b}}$ is significantly higher in the HMXB case (thick, red line), starting to rise from around $z = 20$ due to the heating of the IGM. The global value becomes positive just before $z = 14$, and by the end of the simulation it approaches the high-T_K limit. The evolution of the global 21-cm signal is similar to

that of the analytical and semi-numerical models in the literature (e.g. Pritchard and Furlanetto, 2007; Mesinger et al., 2013), apart from the timing of this transition, which depends on the specific assumptions made about the ionizing and X-ray sources. One difference from these models is that we do not model the early Ly α background, but assume efficient WF coupling at all times. When this assumption is not made, weaker WF coupling early on produces a shallower absorption signal (typically $\overline{\delta T_{\rm b}}_{\rm min} > 180\,{\rm mK}$ instead of $\overline{\delta T_{\rm b}}_{\rm min} \sim 200\,{\rm mK}$ as in our case). While the incomplete Ly α coupling could be an important effect at the earliest times ($z > 20$), we focus on the subsequent X-ray heating epoch, where this effect should have minimal impact.

The lower left-hand panel of Fig. 3.8 shows the rms or standard deviation of the $\delta T_{\rm b}$ as a function of redshift, calculated according to Equation 3.1 and for the resolution specified below Equation 3.3. Before the X-ray heating is able to significantly impact the cold IGM, our two scenarios have a similar rms evolution, which is dominated by the density fluctuations and the adiabatic cooling of the IGM. Later on, the rms drops slightly in the HMXB case due to the local X-ray heating from the first sources, which brings the local IGM temperature closer to that of the CMB. As the characteristic scale of the X-ray heating fluctuations increases (cf. the power spectra evolution), the rms of the X-ray case starts rising again and peaks at ~ 11.5 mK around $z \sim 16.5$. Thereafter, the rms fluctuations gradually decrease due to the $\overline{\delta T_{\rm b}}$ rising towards positive values and the 21-cm absorption turning into emission. The rms in the stellar-only case continuously rises, as the density fluctuations increase. Although local ionization introduces small-scale $\delta T_{\rm b}$ fluctuations, they are smoothed out by the beam- and bandwidth-averaging.

At later times in the HMXB case, the rms asymptotically approaches the high-$T_{\rm K}$ limit, but does not quite reach it by the end of the simulation at $z = 12.7$. In the stellar-only case, the rms of $\delta T_{\rm b}$ is driven purely by the density (and later on to a small extent ionization) fluctuations. The rms, therefore, continues to rise as structure formation continues, ionization begins, and the IGM further cools. Note that in reality the temperature of the neutral IGM is likely to be impacted by shock heating, and we leave this to future work. The features of the rms evolution and their timing are dependent on the resolution available and, hence, on the details of the radio interferometer.

The higher order statistics of $\delta T_{\rm b}$ are also affected by the inclusion of X-rays. The skewness of $\delta T_{\rm b}$ is shown in the top-right panel of Fig. 3.8. In both cases, the skewness starts close to zero, tracking the initial, Gaussian density fields. The skewness of the density field is, in

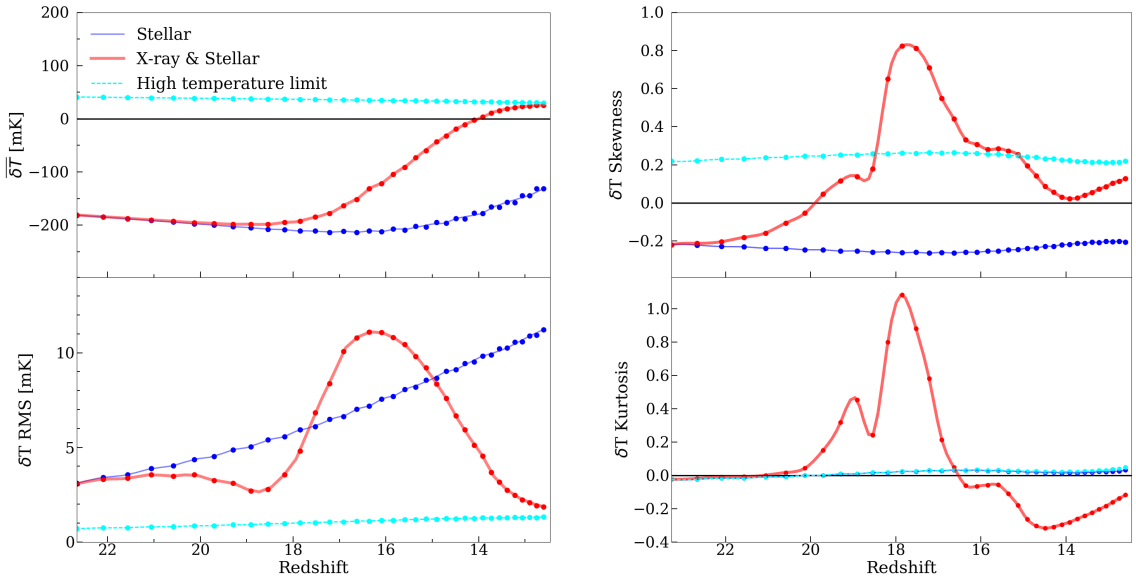

Figure 3.8: Statistics from the 21-cm signal from both our simulations as well as the high-$T_{\rm K}$ limit. The top left panel shows the mean value of $\delta T_{\rm b}$, the bottom left panel the rms, the top-right panel the skewness and the bottom right the kurtosis. The points are the results calculated from smoothed coeval boxes from our simulations and the fitted line is a cubic spline of these data.

fact, not exactly zero as non-linear structure formation introduces positive skewness to the PDF of the density field (Coles and Frenk, 1991) which translates to a negative skewness in the PDF of $\delta T_{\rm b}$ when it is in absorption. The skewness then gradually increases in the HMXB case as hot regions surrounding sources positively skew the data. The skewness from the HMXB case then increases rapidly, as large regions of the IGM are heated. The skewness then peaks around $z \sim 18$ with a value of 0.9. After this peak, the skewness decreases as the heating becomes more homogeneous and reaches zero before increasing again and approaching the high-$T_{\rm K}$ limit (shown as dashed, cyan line). The skewness from the stellar-only case remains negative throughout the simulation and only begins to rise toward the end of the simulation when the ionized fraction becomes non-negligible. Once again, the high-$T_{\rm K}$ limit is not valid for the skewness at the early times considered in this work, except for the very last stages of the HMXB case. The skewness from the high-$T_{\rm K}$ limit and the stellar-only case are mirror images of each other due to the fact that they are both dominated by density fluctuations, as $T_{\rm s} \gg T_{\rm CMB}$ in the high-$T_{\rm K}$ limit and the stellar-only case is dominated by the adiabatic temperature of the universe. The high-$T_{\rm K}$ limit is forced to be in emission; whereas the stellar-only case is in absorption. Our skewness results are in reasonable agreement with the ones for low X-ray efficiency

case '$\log \zeta_{\rm X} = 55$' in Watkinson and Pritchard (2015) (cf. their Fig. 13, but note that they show the dimensional skewness, which is the same as the dimensionless skewness multiplied by the corresponding rms).

The kurtosis of $\delta T_{\rm b}$ from our two simulations is displayed in the bottom right panel of Fig. 3.8. Initially, the kurtosis of both models is close to zero as the Gaussian density fluctuations dominate at these early stages. As long-range heating develops, the kurtosis of the HMXB case increases and peaks at $z \sim 18$ with a value of 1.7. Later, the kurtosis decreases again to negative values, before increasing once more and approaching temperature saturation (shown as dashed, cyan line). This statistic follows roughly the same pattern as the skewness, but with somewhat different functional shape and timing. The stellar-only case kurtosis closely tracks that of the high-$T_{\rm K}$ limit throughout the simulation, deviating only slightly at the end due to the small amount of ionization present. As there is no heating in this case, the signal is mostly Gaussian with small non-Gaussianity only arising from the density fluctuations.

3.7 Discussion

We present the first large-volume, fully numerical structure formation and radiative transfer simulations of the IGM heating during the Cosmic Dawn by the first stellar and X-ray sources. We simulate the multi-frequency transfer of both ionizing and X-ray photons and solve self-consistently for the temperature state of the IGM. While the exact nature and properties of the first X-ray sources are still quite uncertain, our results demonstrate that, under a reasonable set of assumptions, these sources produce significant early and inhomogeneous heating of the neutral IGM and, thus, impact considerably the redshifted 21-cm signals. We focus on these expected 21-cm signals from this epoch and its statistics throughout this paper.

In this work, we consider relatively soft-spectrum X-ray sources, which trace the star formation at high redshift. At these high redshifts, these sources are still fairly rare and for reasonable assumed efficiencies, the addition of X-rays does not affect significantly the evolution of the mean fractions of H II and He II. The fraction of He III, however, is boosted by almost an order of magnitude compared to the stellar-only case, although remaining quite low overall.

The high energies and long mean free paths of the hard X-ray radiation make it the dom-

inant driver of the heating of the neutral IGM. Pop. II stars, even massive ones, do not produce a significant amount of such hard radiation. Therefore, both the morphology and the overall amount of heating change dramatically when X-ray sources are present. The mean and the median temperature both increase considerably compared to the stellar-only case, with the mean eventually reaching $\sim 10^3$ K by $z \sim 13$ (the median, which only reaches ~ 200 K, better reflects the neutral IGM state as it is less sensitive to the very high temperatures in the ionized regions). The X-ray heating is long-range and, therefore, widely distributed throughout the IGM. This heating is also highly inhomogeneous, as evidenced by the temperature PDFs, maps, and evolution seen in the lightcone visualisations. The neutral regions are heated by the X-ray sources and go fully into 21-cm emission with respect to the CMB before $z = 13$, while with stellar-only sources the IGM remains in absorption throughout the Cosmic Dawn. The presence of X-rays, therefore, results in an early, but extended ($\Delta z \sim 7$) transition into emission.

The 21-cm fluctuations initially ($z > 20$) track the density fluctuations due to the still insignificant heating and ionization fluctuations. However, the temperature fluctuations due to X-ray heating quickly boost the large-scale 21-cm fluctuations to much higher values. At a resolution of $\sim 10 - 12$ arcmin for redshifts 15 – 17, the fluctuations are large enough to be a factor of several above the expected noise level of SKA1-Low, which implies the possibility of observing not only power spectra, but also coarse images of the 21-cm signal from the Cosmic Dawn. For the same resolution, the $\delta T_{\rm b}$ rms in the presence of X-rays peaks at ~ 11.5 mK around $z \sim 16.5$. As the X-rays heat the neutral IGM, a broad peak develops at $k \sim 0.1$ comoving Mpc^{-1}, corresponding to spatial scale of about 43 comoving Mpc, at $z \sim 14 - 15$. As the IGM heats up and the absorption gradually turns into emission, the 21-cm fluctuations for the HMXB case decrease and approaches the high-$T_{\rm K}$ limit, which is not fully reached by the end of our simulation ($z \sim 12.7$), even though by that time the mean IGM is heated well above the CMB temperature. In contrast, the stellar-only case fluctuations are still increasing steeply by $z \sim 12$, as they are driven by cold IGM.

In the HMXB case, the distribution of the $\delta T_{\rm b}$ fluctuations shows a clear non-Gaussian signature, with both the skewness and kurtosis peaking when the fluctuations start rising. By the end of the simulation, the skewness and kurtosis approach, but do not reach, the high-$T_{\rm K}$ limit. For soft radiation sources, the non-Gaussianity is driven by density fluctuations only, producing a smooth evolution.

The often-used high spin temperature limit, $T_{\rm S} \gg T_{\rm CMB}$, is not valid throughout the X-ray heating epoch as long as any IGM patches remain cold. When X-rays are present, even after the IGM temperature rises above the CMB everywhere (and thus the 21-cm signal transits into emission), significant temperature fluctuations remain and contribute to the 21-cm signal. The neutral regions do not approach to the high-temperature limit until quite late in our model, at $z \sim 12$. This asymptotic behaviour can readily been seen in the power spectra and statistics of the 21-cm signal. Soft, stellar-only radiation has short mean free paths and, therefore, never penetrates into the neutral regions, leaving a cold IGM.

Previous work in this area has largely been limited to approximate semi-analytical and semi-numerical modelling (e.g. Pritchard and Furlanetto, 2007; Mesinger et al., 2013; Fialkov and Barkana, 2014; Shimabukuro et al., 2015; Watkinson and Pritchard, 2015; Kubota et al., 2016). By their nature, such approaches do not apply detailed, multi-frequency RT, but rely on counting the photons produced in a certain region of space and comparing this to the number of atoms (with some correction for the recombinations occurring). The difference between the two determines the ionization state of that region. The X-ray heating is done by solving the energy equation using integrated, average optical depths and photon fluxes, and often additional approximations are employed as well (e.g. Watkinson and Pritchard, 2015). These methods typically do not take into account non-linear physics, spatially varying gas clumping or absorbers, or Jeans mass filtering of low mass sources. These differences make comparisons with the previous results in detail difficult due to the very different modelling employed and would require further study. Nonetheless, we find some commonalities and some disparities with our results, summarised below.

Our thermal history is similar to that of the relevant cases in Pritchard and Furlanetto (2007) (their Case A) and Watkinson and Pritchard (2015) (their case '$\log \zeta_{\rm X} = 55$'). We find a quite extended transition between 21-cm absorption and emission, from the formation of the first ionizing and X-ray sources at $z \sim 21$ all the way to $z \sim 13$. This transition is somewhat more protracted than the one in the most similar scenarios ($f_{\rm x} = 1$ and 5) considered in Mesinger et al. (2013), likely due to the higher star formation efficiencies assumed in that work.

We find a clear X-ray heating-driven peak in the 21-cm power spectra at $k = 0.1 - 0.2\,{\rm Mpc}^{-1}$, similar to the soft X-ray spectra peak found in Pacucci et al. (2014a) and at similar redshift (z$\sim$ 19 15 $-$ 16; though this depends on the uncertain source efficiencies).

Results from their peak power, at $\Delta_{21\mathrm{cm}} \sim 14\,\mathrm{mK}$, are in rough agreement with our results. The general evolution of the power spectra found in Pritchard and Furlanetto (2007) appears similar, with the fluctuations at $k = 0.1\,\mathrm{Mpc}^{-1}$ also peaking at $z \sim 15-16$ (although that only occurs at $z \sim 12-13$ for the scenario with less X-rays, again suggesting a strong dependence on the source model). The power spectra found are in reasonable agreement with our results, with peak values of $\Delta_{21\mathrm{cm}} \sim 19\,\mathrm{mK}$ or $\Delta_{21\mathrm{cm}} \sim 11.5\,\mathrm{mK}$ depending on the source model used by them, compared to $\Delta_{21\mathrm{cm}} \sim 14\,\mathrm{mK}$ for our HMXB case.

The 21-cm skewness from the X-ray heating epoch is rarely calculated, but Watkinson and Pritchard (2015) recently found a very similar evolution to ours (though shifted to somewhat higher redshifts), with a positive peak roughly coinciding with the initial rise of the 21-cm fluctuations due to the temperature patchiness. Their corresponding 21-cm δT_b PDF distributions during the X-ray heating epoch significantly differ from ours, however. At the epoch when T_S reaches a minimum, the semi-numerical model predicts a long tail of positive δT_b, which does not exist in the full simulations. Around the $T_\mathrm{S} \sim T_\mathrm{CMB}$ epoch, our distribution is quite Gaussian; while Watkinson and Pritchard (2015) find an asymmetric one (though, curiously, with one with close to zero skewness, indicating that skewness alone provides a very incomplete description). Finally, in the $T_\mathrm{S} \gg T_\mathrm{CMB}$ epoch, the two results both yield Gaussian PDFs, but the simulated one is much narrower.

Our models confirm that, for reasonable assumptions about the presence of X-ray sources, there is a period of substantial fluctuations in the 21-cm signal caused by the patchiness of this heating and that this period precedes the one in which fluctuations are mostly caused by patchiness in the ionization. However, since the nature and properties of X-ray sources remain unconstrained by observations, other scenarios in which the heating occurs later, are also allowed. The currently ongoing observational campaigns of both LOFAR and MWA should be able to put constraints on the presence of spin temperature fluctuations for the range $z < 11$, which would then have clear implications for the required efficiency of X-ray heating at those and earlier redshifts. In the future, we will use simulations of the kind presented here to explore other possible scenarios, for example heating caused by rare, bright sources, as well as the impact of spin temperature fluctuations on all aspects of the 21-cm signal, such as redshift space distortions.

Chapter 4

Evaluating the QSO contribution to the 21-cm signal from the Cosmic Dawn

4.1 Contributions of the Authors

The work presented in this chapter has been submitted, see Ross et al. (2018). The contributions of the authors of this paper to the work are as follows:

- **Hannah Ross:** modified C^2-Ray to include QSO sources (see Appendix B), helped develop the mutliphase method (see Appendix C), ran simulations, ran the Lyman-α coupling coefficient semi-numerical code, calculated the statistics, produced all plots for the paper and wrote most of paper.
- **Keri Dixon:** provided the QSO source lists and assisted with paper writing.
- **Raghunath Ghara:** updated and then provided his code to calculate Lyman-α coupling coefficients and assisted with paper writing
- **Ilian Iliev:** helped develop the multiphase method, paper writing, gave general guidance and secured computing time.
- **Garrelt Mellema:** helped develop the multiphase method and assisted with paper writing

4.2 Versions of C^2-Ray used in this Work

This project required a significant amount of additional work in the form of code development. Firstly, in order to simulate QSOs, C^2-Ray had to be updated to be able to handle more than one type of X-ray source (see Appendix B for details).

While the method described in Appendix A was sufficient for that work, it cannot be used to correct the results from simulations extending beyond z≈11 (see Appendix C section C.2.3 for details). Therefore, we updated the algorithm to follow the temperature of the neutral gas without contributions from hot ionized regions. This new multiphase algorithm is presented in Appendix C section C.1. In addition to this, results from a number of test boxes are shown to demonstrate the new algorithm's effectiveness in Appendix C section C.2.

The simulations with HMXBs in this chapter use the original code used in Chapter 3, which is acceptable as temperature saturation is reached before z≈11. The stellar only simulation also used the original code as the temperature is simply taken to be the adiabatic temperature of the Universe. However, simulations with QSOs as the only X-ray sources present were run with the new multiphase method.

4.3 Introduction

This Chapter focuses on the investigation of QSO heating during the CD. Here we extend the suite of simulations presented in Chapter 3 with three additional large scale fully numerical RT simulations, now including QSO sources. We also investigate the impact of an inhomogeneous Lyman-α background and its possible impact on the 21-cm signal in combination with fluctuations from long range heating.

The density fields and source catalogues are once more obtained from CUBEP^{3}M (see section 2.1 for further details). The RT simulations were run with a newly updated version of C^2-Ray, see Appendix C. Here QSOs are considered in addition to the original HMXBs investigated in the previous Chapter. The sources are briefly summarized below:

- *Stellar sources:* These sources have luminosities proportional to their host halo's mass and have black body spectra, producing UV photons. These are the sources driving reionization in our simulations and do not contribute to heating.
- *HMXBs:* Like the stellar sources these sources have luminosities proportional to

Spectral index	S1	S2	S3	S4	S5
$\alpha_{\rm x}^{\rm h}$	–	–	–	-1.5	-1.5
$\alpha_{\rm x}^{\rm q}$	–	-0.8	-1.6	–	-0.8

Table 4.1: Table showing the spectral index of our X-ray sources in different runs. All simulations include stellar sources with a blackbody spectrum with a temperature 5 $\times 10^4$ K. $\alpha_{\rm x}^{\rm h}$ and $\alpha_{\rm x}^{\rm q}$ are the spectral indices of the HMXB and QSOs respectively.

their host halo's mass but have a power-law spectra: $L_{\rm h}(E) \propto E^{-\alpha_{\rm x}^{\rm h}}$. They produce only X-ray photons.

- *QSOs:* These sources have luminosities calculated from an extrapolation of the QLF. Like HMXBs they also have a power-law spectra but with a different index, $L_{\rm q}(E) \propto E^{\alpha_{\rm x}^{\rm q}}$. Two QSO models with separate values of $\alpha_{\rm x}^{\rm q}$ were considered (described in detail in section 2.2.3).

The parameters for the RT simulations run are outlined in Table 4.1. Spectral indices $\alpha_{\rm x}^{\rm h}$ and $\alpha_{\rm x}^{\rm q}$ are given in terms of energy.

The amount of emitted soft (i.e. non-ionizing) UV photons that redshift into the Lyman-α resonance in the CD is highly uncertain as we know very little about the sources that created them (most likely Pop III stars). Here we consider the two extremes: very early and very late Lyman-α saturation scenarios in order to demonstrate the full range of possible impact of these uncertain source parameters. We leave the more detailed analysis of the impact of Ly-α photons for future work.

Early minihalos (halos with masses below $\sim 10^8 {\rm M}_\odot$ hosting Pop III stars) may contribute significantly to the Lyman-α background, which could result in Lyman-α saturation occurring earlier than expected. In the most extreme scenario a strong Lyman-α background may have already been built up by minihalos before our simulation has begun ($z \sim 23$), as was assumed in Chapter 3. This scenario is hereafter referred to as *early Lyman-α saturation*.

On the other end of the possible range, mini-halos might contribute very little Lyman-α radiation, so the Lyman-α background is built up by the HMACHs and LMACHs alone. In this scenario we must consider the inhomogeneous background they produce using the updated method from Ghara et al. (2016) described in section 2.4.1. This scenario is referred to as *late Lyman-α saturation* throughout this Chapter.

4.4 Evolution of the signal

In Fig. 4.1, we show the $\delta T_{\rm b}$ lightcones, slices from the position-redshift image cube, for each of the five simulations. In this figure, we only show late Lyman-α saturation, meaning that all models start with $\delta T_{\rm b} = 0$. When compared to the results in Paper I, we see that the inclusion of the Lyman-α coupling effect in this case yields a weaker absorption signal, since heating is already under way before Lyman-α saturation ($y_\alpha \gg 1$) is reached. The inhomogeneous Lyman-α background also softens the features present before Lyman-α saturation (z$\approx$17). Unless stated otherwise, the late (z$\approx$17) Lyman-α saturation case is the default.

After the Lyman-α background saturates, $\delta T_{\rm b}$ in simulation S1 remains in absorption for the rest of the simulation due to the lack of photons with long enough mean free paths to penetrate and heat the neutral IGM. In the absence of X-ray heating the IGM cools adiabatically as the universe expands. Hence, the transition from absorption to emission never occurs in this case. At later times, ionized patches indicating the beginning of significant ionization can be seen as holes ($\delta T_{\rm b} \approx 0$) in the HI distribution.

In S2 and S3, the signal remains largely in absorption for longer than in S4 and S5, due to the total photon budget from the QSO sources being much lower than that of the HMXBs. However, at z $\approx$ 16.5, patches with a higher $\delta T_{\rm b}$ start to develop around the QSO sources as they locally heat their surroundings. The sizes of these heating regions depend strongly on the spectra of the QSOs, with the harder-spectrum simulation S2 yielding noticeably more widespread heating. These regions are initially rare, but as more QSO sources form, they begin to overlap and bring the higher density regions into emission. Eventually, the QSOs become sufficiently numerous to heat the voids, and the entire simulation volume transitions into emission at z$\sim$9.5 and 9 for S2 and S3, respectively. Heating progresses more rapidly in S2, again due to the spectrum of the QSOs being harder than those in S3. In these models, temperature saturation ($T_{\rm S} \gg T_{\rm CMB}$) is not reached until reionization is underway. Additionally, early ionized bubbles are less sharply outlined, since the $\delta T_{\rm b}$ is closer to zero than it is in simulations S1, S4, and S5.

In S4 and S5, the long-range X-ray heating due to HMXBs produces an earlier and less extended transition from absorption into emission at z$\sim$ 14.5 when compared to the QSO-only models (S2 and S3). Before temperature saturation is reached (z $\approx$ 13.5), large-scale fluctuations throughout both S4 and S5 can be seen. These initial emission regions expand quickly, some increasing to tens of Mpc in size by $z \sim 15.5$. Additional heated regions are

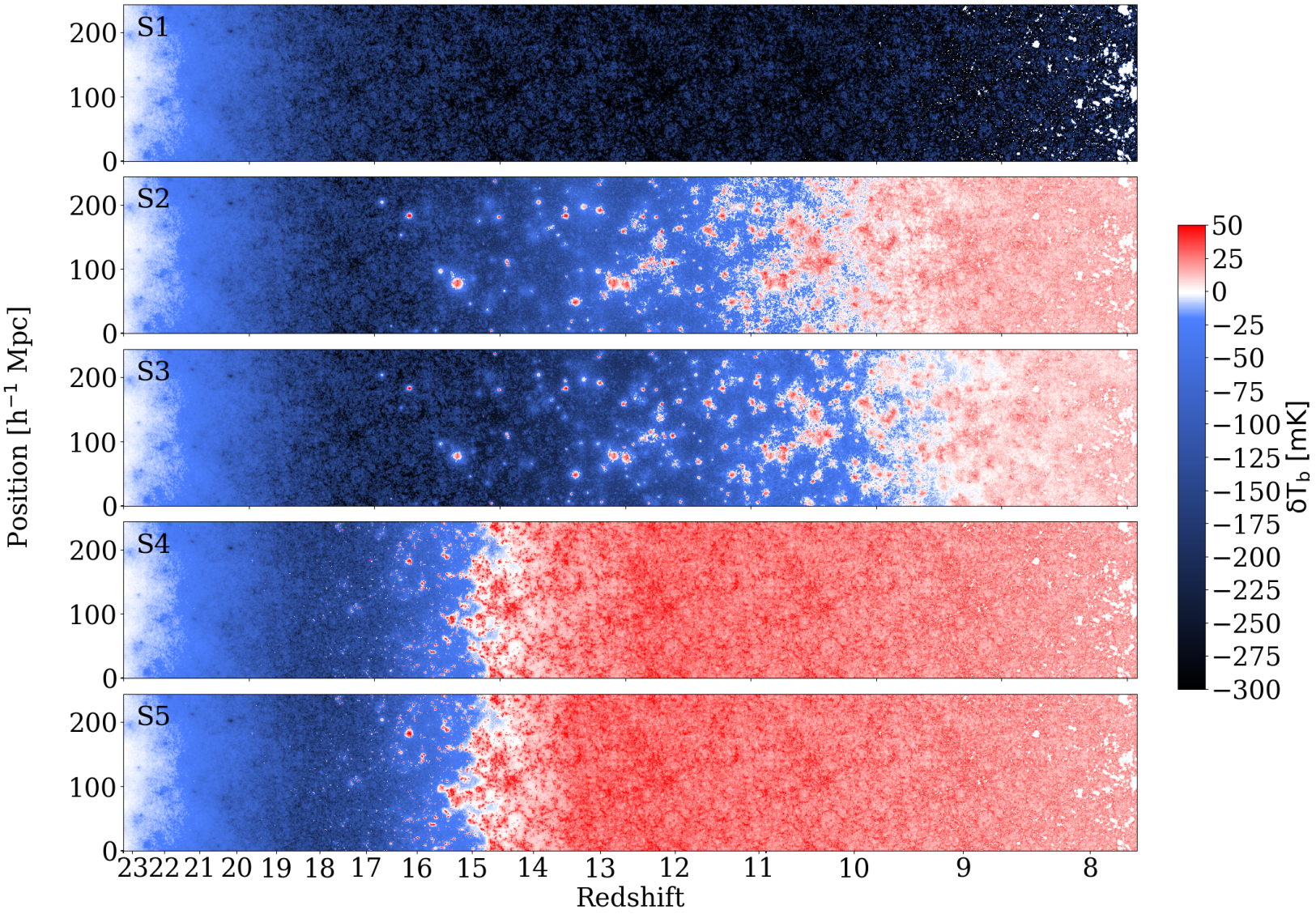

Figure 4.1: The position-redshift lightcone of $\delta T_{\rm b}$ from the five different simulations considered in this study. Here, we can see the different geometries, evolutions, and timings produced by the different source models. The details of the source models are given in Table 4.1. These lightcones are for the case of late Lyman-α saturation.

z	8.95	15.98	25.25
θ [arc min]	6.0	10.3	15.8
$\delta T_{\rm b}$ [mK]	4	5	20

Table 4.2: Table showing the expected noise on SKA1-Low from (Koopmans et al., 2015) for a maximum baseline of 1.2 km and an integration time of 1000 hrs.

visible in S5 at z $\approx$ 16 where QSOs have formed; however, the transition from absorption to emission and temperature saturation happens at roughly the same time in these two models, as the heating from HMXBs dominates over that from the rare QSOs.

In Fig. 4.2, smoothed mean-subtracted maps of the 21-cm signal are displayed. To generate these maps coeval cubes are smoothed in the two angular directions with a two-dimensional Gaussian beam with a FWHM corresponding to a 1.2 km baseline at the frequency corresponding to the redshift of interest ($\theta_{\rm FWHM} = 0.221(1+z)/1200$). This maximum baseline length approximately corresponds to the planned size of the core of SKA1-Low. Along the line-of-sight (frequency) direction, the data are smoothed with a top-hat function, the width of which is equal to the comoving distance corresponding to $\theta_{\rm FWHM}$. We will use this redshift-dependent resolution throughout this paper. Table 4.2 lists the angular resolution together with the expected rms noise value at this resolution for three representative redshifts.

Fig. 4.2 shows that initially ($z \sim$20) the fluctuations are dominated by fluctuations in the Lyman-α coupling, making the X-ray source models indistinguishable. These fluctuations have a magnitude of just over 10 mK, which is above the value of the expected noise per resolution element for 1000 hrs of integration with SKA1-Low, $\approx$15 mK (Koopmans et al., 2015), at these redshifts, implying that these phases cannot be studied with tomographic images.

However, at z$\sim$16, the SKA1-Low noise levels for deep integrations and our standard resolution are expected to be around 5 mK. All our simulations show fluctuations that exceed this. S4 and S5 exhibit fluctuations around 20 mK, and S2 and S3 have even higher ones, reaching over 40 mK in magnitude. These levels imply that each of these X-ray source models could possibly be imaged directly with SKA1-Low. All models can be distinguished visually at z$\sim$16. Models including HMXBs, S4 and S5, then approach temperature saturation and become indistinguishable from each other after z$\sim$15. Below $z \approx 13$, the low contrast in the S4 and S5 images will make it more challenging to image

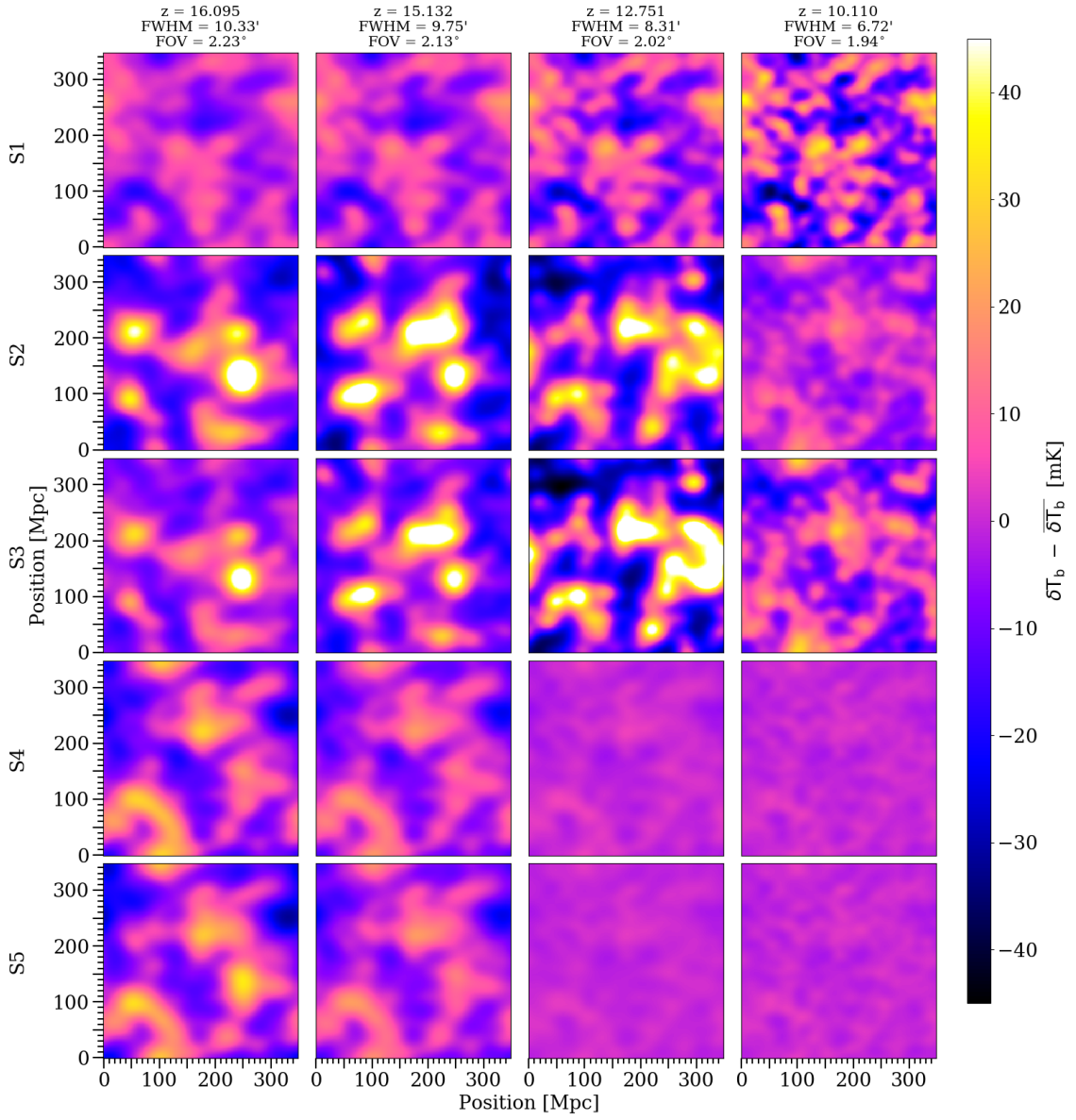

Figure 4.2: Mean-subtracted $\delta T_{\rm b}$ maps, smoothed with a Gaussian beam with a FWHM corresponding to a 1.2 km maximum baseline at the relevant frequency and bandwidth-smoothed with a top hat function (width equal to the distance corresponding to the beam width). The rows correspond to our five models. The columns represent higher to lower redshift from left to right.

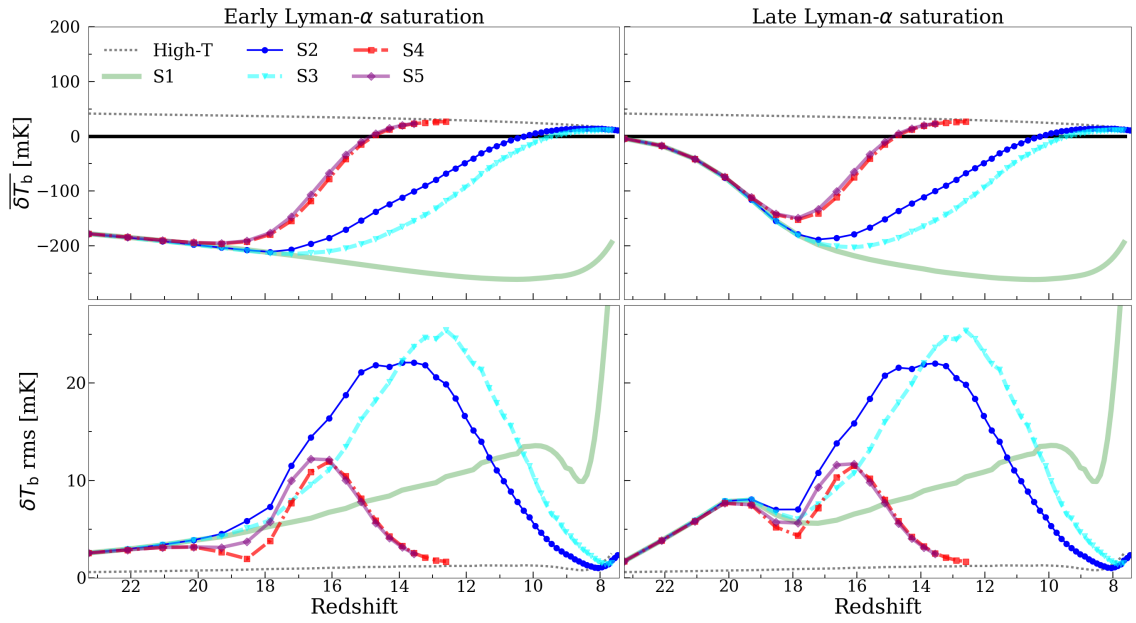

Figure 4.3: The mean and rms of $\delta T_{\rm b}$ for both early Lyman-α saturation (left-hand side panels) and late Lyman-α saturation (right-hand side panels). The high-$T_{\rm S}$ limit is shown to illustrate when temperature saturation occurs in each model.

these stages. The fluctuations in S2 and S3 continue to increase, exceeding 40 mK. The fluctuations in S2 peak at approximately z$\sim$15 and S3 at z$\sim$12. For lower redshifts ($z \approx$ 10), the contrast in the images drops, but may still be enough to allow imaging.

In Fig. 4.3, we show the evolution of the mean $\delta T_{\rm b}$ (top panels) and its rms fluctuations, smoothed to the expected SKA resolution (bottom panels) for both early (left panels) and late (right panels) Lyman-α saturation. In all cases, we also include the high-$T_{\rm S}$ limit case, indicating when the temperature saturation limit is being approached.

When we assume Lyman-α saturation (top left panel), all simulations start with a negative value due to the initially very cold IGM and then drop further as the expanding universe adiabatically cools, until X-ray heating reverses this trend. The timing of this heating depends on the specific X-ray source model. In S1, the IGM remains cold and thus $\overline{\delta T_{\rm b}}$ decreases all the way to $z \sim 11$ when the highest density peaks start to become ionized and cause the signal to increase slightly. The signal in S2 and S3 initially follows a similar pattern, however, after redshift z$\approx$17 heating from the QSOs begins to have an impact and gradually raises the temperature of the neutral IGM until it eventually converges to the high temperature limit at z$\approx$8-9. As could be expected, the harder QSOs in S2 contribute more to heating and produce a weaker absorption signal and earlier transition to emission than the softer ones in S3. The presence of HMXBs has a much greater impact on the

global mean than the QSOs. Consequently, the mean value from S4 begins to increase markedly earlier, $z \approx 20$, in $\overline{\delta T_{\rm b}}$ and reaches a positive value already around $z \approx 14$, and approaches temperature saturation at $z \approx 13$.

When the inhomogeneous Lyman-α background is taken into account (top right panel), $\overline{\delta T_{\rm b}}$ instead starts at zero, as the Lyman-α coupling is inefficient at these early times. This mean value then decreases to meet the fully coupled cases, as the Lyman-α background appreciates. Late Lyman-α coupling considerably reduces the length of the period of strong absorption compared to the case of early Lyman-α saturation. The absorption signal peak magnitude is noticeably reduced in the early-heating cases S4 and S5, but is unaffected in the late-heating ones since then Lyman-α saturation is achieved before the absorption signal peaks.

The lower panels of Fig. 4.3 show the rms (i.e the standard deviation) of $\delta T_{\rm b}$, smoothed to the expected resolution of SKA. The lower left-hand panel of Fig. 4.3 displays the rms calculated for the case of early Lyman-α saturation. Before the X-ray heating is able to have significant impact on the cold IGM, all scenarios follow a similar rms evolution, which is dominated by the density fluctuations and the adiabatic cooling of the IGM. The fluctuations in S1 continue to be driven by density fluctuations, which increase as structure formation progresses. At z$\sim$10, ionized bubbles begin to grow around the sources, increasing $\delta T_{\rm b}$ in dense regions near sources. As the deep absorption signal dense regions become weaker the rms decreases. As these bubbles continue to expand $\delta T_{\rm b}$ approaches zero in increasingly large regions around the sources, contrasting with the signal from the neutral IGM and increasing the rms. The high-$T_{\rm S}$ limit follows the same evolution, but the rms fluctuations are much lower due to the signal being assumed to be in emission rather than absorption.

The rareness of QSOs in S2 and S3 introduces large-scale heating fluctuations, which increase the peak rms values in by a factor of about two compared to S4 and S5, before the rms decreases in all heating scenarios as heating saturation is approached. The softer QSO model S3 peaks somewhat later than S2 as softer X-ray photons heat the local IGM which rapidly reaches temperature saturation. The additional heating fluctuations due to the QSOs yield a slightly earlier rms fluctuations peak in S5 (z$\approx$ 16.5) compared to S4 (z$\approx$ 16).

In all simulations, the inclusion of the inhomogeneous Lyman-α (Fig. 4.3, bottom right) background boosts the early fluctuations, which results in an additional peak in the rms

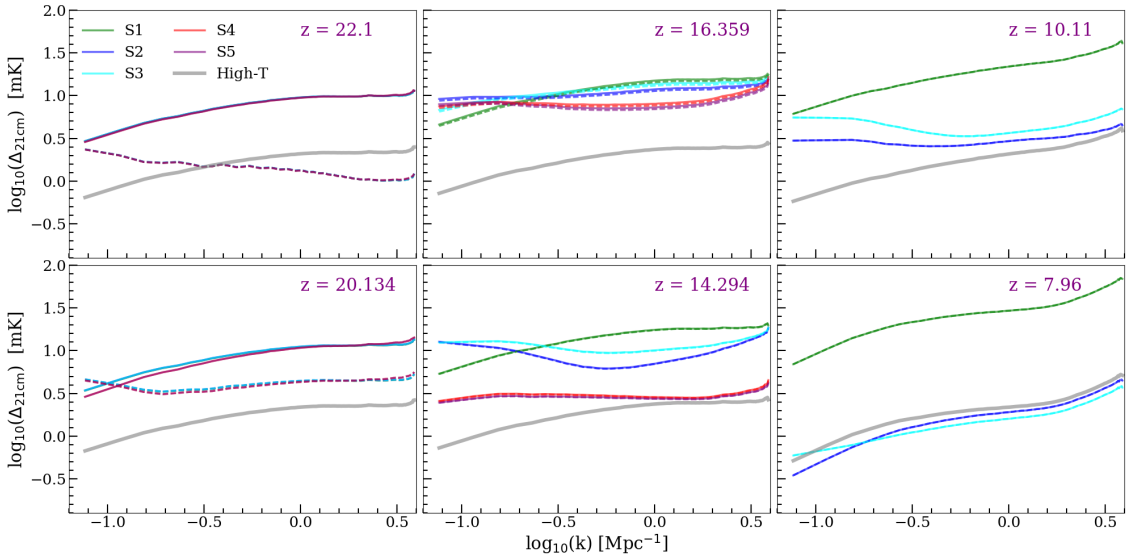

Figure 4.4: The 21-cm power spectra from our simulations at several key stages of the evolution assuming early (solid lines) and late (dashed lines) Lyman-α saturation. S4 and S5 are not shown once they have reached temperature saturation as they are equivalent to the high-$T_{\rm S}$ limit. For reference, the high-$T_{\rm S}$ limit is also indicated, as labelled.

at z$\sim$20. These additional fluctuations come from inhomogeneities in the Lyman-α background being introduced into the signal. Later-time fluctuations ($z < 18$) are unaffected since by then the Lyman-α background saturates.

4.5 Power Spectra

In Fig. 4.4, we show the power spectra of $\delta T_{\rm b}$ at several key redshifts. At the beginning of the simulation ($z \sim 22$), all simulations give identical results with very high (absorption) signal for early Lyman-α saturation and very low signal for late Lyman-α coupling, since in the latter case, the 21-cm signal is still coupled to $T_{\rm CMB}$. At this stage, the heating has not yet had a significant impact. Thus in the early Ly-α saturation cases, the 21-cm fluctuations simply follow the density ones, with a boost due to the strong absorption compared to the high-$T_{\rm S}$ limit. If the Lyman-α background has not yet saturated, the power is suppressed on all scales. There is more power on the large scales, as the Lyman-α couples the regions close to the stars first and the large distances between these regions boost the power on large scales.

By z $\sim$20, Lyman-α coupling has become efficient in the late Lyman-α saturation case. The power has increased on all scales compared to the beginning of the simulation, par-

ticularly the very large scales. There is more power at large scales in the late Lyman-α saturation case than the early Lyman-α saturation case due to the large-scale fluctuations introduced by the inhomogeneous Lyman-α background. In the case of late Lyman-α saturation, the models remain almost the same, with only a slight difference between the models with and without HMXBs.

By $z \approx 16$, the Lyman-α background is sufficiently built up so that the early and late Lyman-α saturation results converge. The small-scale power in simulations S4 and S5 begins to decrease, as X-ray heating washes out small-scale temperature fluctuations; whereas on larger scales, the power is slightly boosted for the same reason. S2 and S3 follow a similar pattern, but have more power on large scales due to the rareness of the brighter QSOs, producing a strong contrast between a few strongly heated regions and other areas still in absorption. On smaller scales, they also have more power, as significant heating has not extended to much of the simulation and hence has not washed out the small-scale temperature fluctuations from previous QSO activity.

By $z \approx 14$, the 21-cm power is starting to decline in S2 and S3 due to large-scale heating fluctuations introducing strong contrast between the hot and cold regions. The power is declining more rapidly in S4 and S5, as the transition to emission takes place convergence to the high–$T_{\rm S}$ limit (which can then be used to describe the rest of the EoR correctly) is approached. The power in S2 and S3 is still boosted on large scales compared to that of S1, but long range heating is beginning to suppress the power on smaller scales.

At $z \approx 10$ S4 and S5 have reached temperature saturation, while S2 and S3 approach temperature saturation much more gradually and later ($z \approx 7.9$), with slightly more power remaining in large scale fluctuations. S3 evolves more slowly than S2 due to the QSOs having softer spectra, thus heating more locally. Power on all scales in S2 and S3 is lower than in the high-$T_{\rm S}$ case as reionization has begun. Regions of the simulation that have not yet reached temperature saturation have values closer to the zero-signal coming from ionized regions, which decreases the magnitude of the fluctuations. In the fully saturated case there is more contrast between the emission signal from heated, neutral regions and zero signal from ionized regions, so the magnitude of fluctuations is greater.

The redshift evolution of several k-modes is shown in Fig. 4.5. Panels on the left-hand side show results from the early Lyman-α heating. On large scales, X-ray heating from HXMBs initially suppresses fluctuations, as heating weakens the absorption signal from the densest regions around sources. As these regions move towards emission, the large-

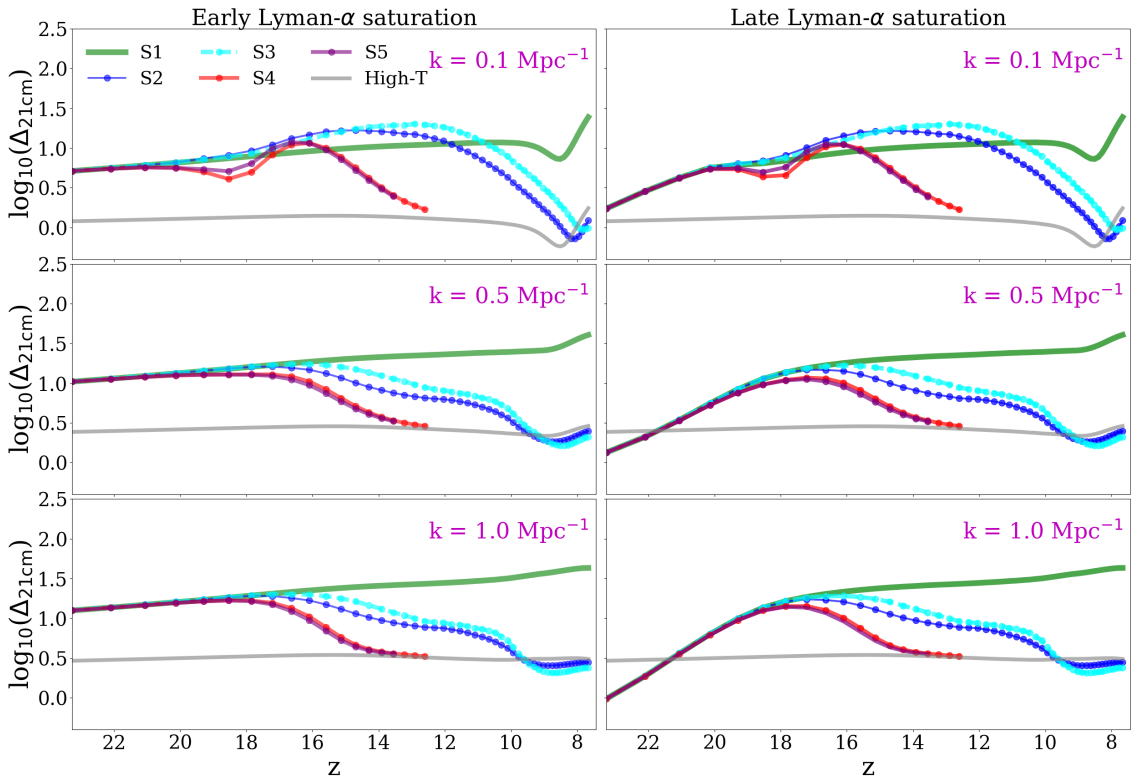

Figure 4.5: The evolution of the 21-cm power spectra modes at $k = 0.1, 0.5$, and $1\,\mathrm{Mpc}^{-1}$ for all X-ray source models. On the left-hand side panels, power spectra modes from the early Lyman-α saturation scenario are displayed, and on the right, the late Lyman-α saturation. The high-T_S limit is displayed to illustrate temperature saturation.

scale fluctuations are boosted before decreasing once more, as temperature saturation approaches. When in combination with HMXBs, QSOs follow a similar pattern. However, with QSOs alone, there is no trough present as many high density regions remain in strong absorption due to the rarity of the QSOs.

The evolution on smaller scales is somewhat similar, however, at no point are the fluctuations boosted by X-ray heating. The removal of deep adsorption signals from dense regions around sources causes this decrease in power, occurring more rapidly in models with HMXBs as X-rays are emitted from each halo. Power on small scales is suppressed more rapidly in S2 as long range heating is more significant.

Fluctuations including the inhomogeneous Lyman-α background from late Lyman-α saturation are shown in the right-hand side panels. Initially, fluctuations are suppressed in all models until the Lyman-α background has been established. This Lyman-α case particularly impacts the fluctuations in S4 and S5, decreasing the peak value at $z = 16$ to $\log_{10}(\Delta_{21\mathrm{cm}}) \approx 0.9$. The peak values of S2 and S3 occur later, after Lyman-α saturation

is reached and so are unaffected.

4.6 Non-Gaussianity of the 21-cm signal

The power spectra alone cannot be used to fully describe $\delta T_{\rm b}$ fluctuations from X-ray heating as they are highly non-Gaussian. Therefore, we consider the higher moments (skewness and kurtosis) of the one-point statistics of the 21-cm signal produced from our simulations. We use the same dimensionless definitions for skewness and kurtosis used in Chapter 3

These quantities are all smoothed to the resolution of SKA1-Low and are calculated from coeval simulation boxes. The fact that QSO sources are stochastic and do not trace the Gaussian density distribution leads to a dramatic increase in non-Gaussianity. Therefore, compared to the other measures discussed, these higher order statistics demonstrate the most extreme differences between models with and without QSOs.

In Fig. 4.6, we show the evolution of the skewness (top) and kurtosis (bottom) for both early (left) and late (right) Lyman-α saturation. Simulation S1 shows a flat, featureless evolution throughout the CD, since the density field at early times is close to Gaussian. The skewness only increases once significant reionization begins at z$\sim$10, when stars ionize the high density peaks. These ionized regions introduce non-Gaussianity into the signal, as shown in previous studies (Mellema et al., 2006b).

In all cases with X-ray heating, the skewness initially follows that of S1 until $z \sim$20 when QSOs begin to form in S2, S3, and S5. In the hard-spectra QSO cases S2 and S5, the skewness increases rapidly; while in the softer-spectrum case S3, it increases more gradually. Similarly in S4, the skewness increases gradually (and peaks at a lower value), since non-Gaussianities are added by the (softer-spectrum) HMXBs. The maximal skewness from S5 is 4.5, or over four times greater than the value (~ 1) obtained from S4. The peak in S5 also occurs somewhat earlier, at $z \sim 18.5$ rather than $z \sim$18. The maximal skewness values obtained from S2 (2.7) and S2 (2.0), respectively, are intermediate between those from S4 and S5, while still much higher (by factors of 27 and 20, respectively) than the value found for S1 (0.1). The largest value for S2 occurs earlier (at $z \sim$17.5) than that of S3 (z$\sim$16), as the harder QSOs heat more rapidly the cold IGM patches responsible for driving the 21-cm signal fluctuations.

The inhomogeneous Lyman-α background (Fig. 4.6, right panels) introduces additional

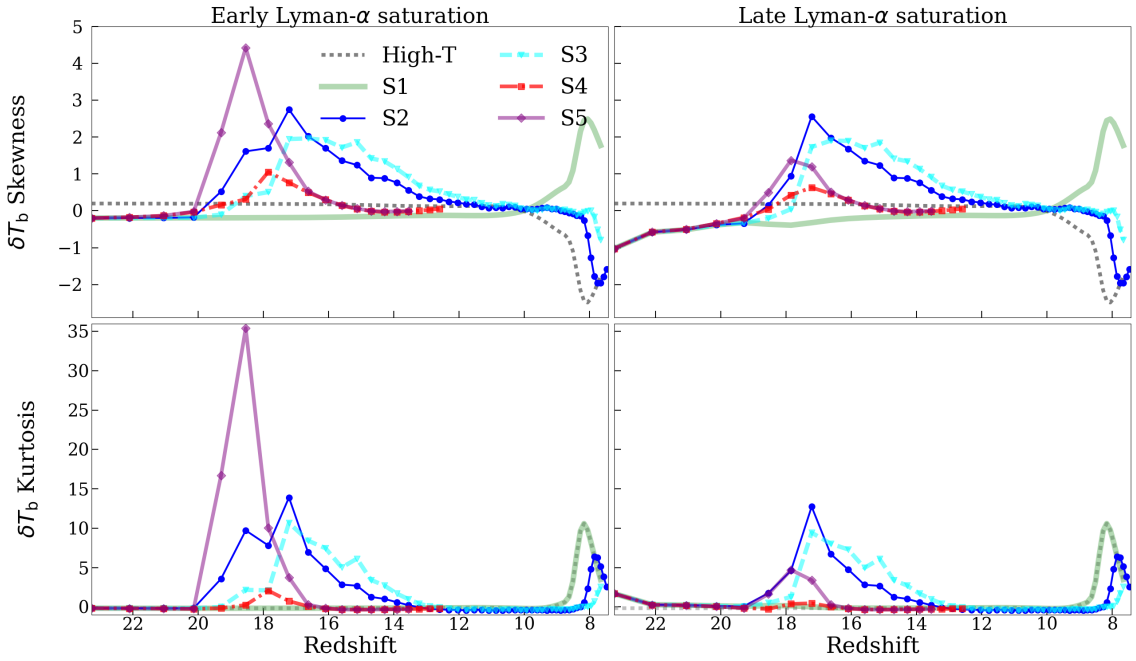

Figure 4.6: The evolution of the skewness (top panels) and kurtosis (bottom panels) of $\delta T_{\rm b}$ for all simulations, as labelled, for early Lyman-α saturation (left panels) and late one (right panels).

non-Gaussian fluctuations of $\delta T_{\rm b}$ due to regions around the sources that first become decoupled from $T_{\rm CMB}$. These dominate over the density fluctuations and give rise to the negative skewness observed in all cases, which then gradually increases as the Lyman-α background builds up as more of the simulation volume decouples from $T_{\rm CMB}$. Thereafter ($z < 19$), heating starts, and temperature variations begin to impact the signal. In all cases, the skewness increases and eventually peaks; however in the X-ray source models where heating occurs earlier, the peaks due to heating fluctuations are suppressed compared to the early Lyman-α saturation. This suppression is particularly evident for simulation S5.

The kurtosis (Fig. 4.6, low panels) follows the same qualitative pattern, but with an even more extreme difference between the QSO and non-QSO cases. The maximal value of the kurtosis from S4 (~ 1) is 17 times smaller than the corresponding value from S5. The kurtosis yields a more notable difference between the two QSO models, with S2 reaching a peak value of 14 vs. peak of just 10 for S3 (both are significantly greater than the value from S1, which is close to zero). The greater amount of heating from the harder QSOs in S2 again leads to a larger deviation from Gaussian fluctuations observed.

When Lyman-α coupling is not complete before heating begins (Fig. 4.6, bottom right), the

maximal kurtosis is also suppressed. This effect is most notable for the X-ray source models S4 and S5, where heating occurs earlier, with the peak in S4 being totally suppressed. The maximum value of S5 is decreased by a factor of 10. The kurtosis of S2 and S3 are slightly suppressed initially but the largest values remain the same as they occur at lower redshift when Lyman-α is closer to saturation.

4.7 Discussion

In this paper, we extended our suite of large-volume, fully numerical RT simulations of X-ray heating of the IGM during the Cosmic Dawn. We introduced two types of QSO sources, with power law spectra slopes of -0.8 and -1.6, and compare their impact to the effects of HMXB X-ray sources (considered previously in Paper I of this series), as well as a new case combining QSO and HMXB sources and a fiducial, stars-only simulation. Unlike HMXBs, the QSOs are rare with a random spatial distribution and luminosities that are not proportional to the host halo mass, though these sources only live in relatively biased HMACHs. The luminosities are instead assigned to match the high-redshift extrapolation of an empirical QXLF. As the precise nature and properties of these early sources remain uncertain, we have chosen an optimistic QXLF that predicts fairly numerous QSOs in order to examine their possible impact. Our simulations show QSO sources may be able to contribute non-trivially to early X-ray heating and also suggest that it is possible to distinguish between soft- and hard-spectra models using the resulting 21-cm signal, particularly via the non-Gaussianity of the signal.

These QSO sources contribute many fewer photons to the X-ray heating than HMXBs, so their overall energy contribution is sub-dominant compared to HMXB sources when both source types are present. On their own, both QSOs models yield a considerably more extended transition of the 21-cm signal from absorption to emission, and spin temperature saturation of the neutral IGM is not reached until reionization itself is under way. The late temperature saturation in the simulations with only QSOs cause the non-Gaussianities of the 21-cm signal due to the appearance of larger ionized regions ($z < 10$) to be lower than in the case of full saturation of the spin temperature. This effect is more pronounced for the QSOs with softer X-ray spectra.

During the CD, heating from QSOs has a more notable impact on the heating fluctuations than on the mean value of $\delta \mathrm{T_b}$. When compared to the stellar only case, the $\delta \mathrm{T_b}$ power

spectrum from all X-ray models show more power on larger scales and less on smaller scales around $z \sim 16$. After this the power spectrum of simulations including HMXBs decreases on all scales as temperature saturation is approached, but the power spectra for the QSO cases experience a further boost on large scales. The rms fluctuations for all X-ray source models are above the expected noise levels for observations with the SKA1-Low core, implying that low resolution tomographic imaging may be possible during the CD. The rare QSOs boost the rms, particularly when they are the only sources of X-ray heating. In this case, the peak value of the fluctuations are ∼10 mK higher than in the HMXB cases, and this peak occurs at lower redshift.

By the far the clearest signature of the QSOs is found in one-point higher order statistics of the 21-cm signal PDF distribution: skewness and kurtosis. An increase in both quantities can be seen both when QSOs are the sole sources driving X-ray heating and when they are present with HMXBs. These strong non-Gaussianities are driven by the rareness of the QSOs introducing fluctuations in the signal largely unrelated to the underlying (mostly Gaussian) density field. However, this increase in non-Gaussianity can be partly suppressed by late Lyman-α saturation, so while an extremely non-Gaussian signal from the CD could indicate the presence of QSOs, a more Gaussian signal would not rule them out.

As well as suppressing the non-Gaussianities, the Lyman-α background fluctuations (in all models) produced by late Lyman-α saturation causes an additional peak in δT_b fluctuations, as has been found in numerous previous works (e.g. Ghara et al., 2015; Watkinson and Pritchard, 2015), which the power spectra show to occur mainly on larger scales. The peak magnitude of these rms fluctuations driven by the Lyman-α background is ≈8 mK, which is well below the expected image noise of 20 mK for imaging with the SKA1-Low core. However, a power spectrum detection of these fluctuations will be possible.

The contribution to long-range heating from our QSO models is larger than that predicted in some other works. For example, Eide et al. (2018) find a negligible contribution to long-range heating from their QSO (referred in their paper as BH) source model. This difference is model dependent, as concluded in Eide et al. (2018), and is in part down to their seeding algorithm only allowing the formation of black holes in haloes greater than $10^{10} M_\odot$, meaning they have a much lower number density of QSOs. This lower density combined with the assumption that QSOs have optically thick, geometrically thin disks (as described in Shakura and Sunyaev (1973)) leads to a somewhat more conservative

heating prediction than our more optimistic models.

Contrarily, the predictions from Yajima and Li (2014) are more in agreement with our findings. They conclude that their single QSO source may be detectable by SKA due to the large emission region it produces. Yajima and Li (2014) follow a prescription similar to Eide et al. (2018) by assuming that black holes accrete at their Eddington luminosities. This discrepancy illustrates that, regardless of the method, the results are highly dependent on the details of the QSO model adopted. Datta et al. (2016) also predict that their individual, bright QSO would be detectable in 1000 hrs integrations with SKA1-Low - also in agreement with our own predictions. Datta et al. (2016) use observations of the low-redshift QSOs to determine their luminosities, a method more comparable to our own.

A key implication of the results presented in this work is that all the X-ray source models investigated go through phases in which the fluctuations in the 21-cm signal are above the expected noise for observations with the core of SKA1-Low (Koopmans et al., 2015), suggesting that at least part of this epoch could not only be studied with power spectra, but could also be imaged. We already concluded this in Paper I for our HMXB case, and we can now extend this conclusion to the cases with rare QSO-like sources. With higher levels of galactic foregrounds and stronger ionospheric effects, imaging around $z \sim 16$ will not be easy, but images of the CD would open the door to a multitude of analysis techniques to extract information from the signal about the astrophysics of the CD. Examples include parameter estimation through deep learning of images (Gillet et al., 2018), the bispectrum (Shimabukuro et al., 2016; Watkinson et al., 2018), and size distributions of features (Giri et al., 2018).

Ghara et al. (2015) and Baek et al. (2010) find power spectra with similar magnitudes to our own. Detailed comparisons to these works are complicated by the fact that (Baek et al., 2010) has no subgrid model and (Ghara et al., 2015) use a subgrid model very different from the one employed in this work. The subgrid modelling is particularly important for comparisons as lower resolutions of these simulations mean that resolved halos do not appear until later in the Cosmic Dawn. Pritchard and Furlanetto (2007) and Pacucci et al. (2014b) also find power spectra with comparable magnitudes. The large scale fluctuations in these works (at k$\approx$ 0.1Mpc^{-1}) peak at the same time as models including HMXBs in this work (z$\approx$16).

Baek et al. (2010) and Watkinson and Pritchard (2015) include measures of the power

spectra and higher order statistics of the 21-cm signal produced by HMXB sources. The peak values for the skewness during the CD found from models S4 and S1 in Baek et al. (2010) are in agreement with the values found in our HMXB model (S4), although the skewness peaks slightly later in their case as z≈11.5. Their skewness peaks later and has a value significantly lower than the values found from our simulations including QSOs. The lack of correlation between our QSO luminosities and their dark matter haloes, along with their rareness are the factors driving this. The difference is far more pronounced in the early Lyman-α saturation case, but still noticeable in the late Lyman-α saturation case. Similarly, the skewness found for all models in Watkinson and Pritchard (2015) have peak values far lower than that of our QSO models but are in in agreement with our HMXB models. The skewness in these models peaked somewhat earlier than the skewness from our models at z≈20. The differences in timing between Baek et al. (2010), Watkinson and Pritchard (2015) and our own models illustrate how sensitive timings are to the properties of the souce models used.

The high level of non-Gaussianities produced by QSOs imply that higher order statistics may be a more useful probe than the power spectra to distinguish between certain source models in the CD. In particular, high non-Gaussianities likely indicate the presence of rare sources such as QSOs. The skewness shows a clear difference between our models including QSOs and our models containing HMXBs as well as models from other works.. Clearly our QSO-like sources introduce a significantly greater amount of non-Gaussianity in the signal. This result further motivates the use of alternative analysis techniques to interpret the signal from the CD to probe for rare X-ray sources.

Conversely, investigating the non-Gaussianities during the EoR could be less insightful than previously hoped if late heating occurs. Despite the additional fluctuations introduced to the kinetic temperature of the gas, non-Gaussianities during the EoR are in fact lower in our QSO-only models. This decrease is due to the magnitude of $\delta T_{\rm b}$ from the neutral regions being lower in the late heating case than for the saturated spin temperature case, removing the brightest points from the signal.

There are a number of potential X-ray sources that may have contributed to the fluctuations in the 21-cm signal in the CD that we have not yet been considered, in particular diffuse gas heated by supernovae-driven winds (e.g. Yajima and Khochfar, 2015). In addition, there is still a large parameter space associated with our current sources, for example varying the star formation efficiency of haloes and hence the luminosity of our HMXBs.

Due to the computational expense of our simulations, we cannot fully explore the huge parameter space associated with the CD. However, the alternative of exploring the parameter space with fast semi-numerical models may well not include all the relevant physics. Ideally, these two approaches should be combined in order to achieve this.

Another limitation of the current work is the modelling of the Lyman-α background. We have included the two most extreme cases, one where very early sources build up a Lyman-α background before the simulation begins (early Lyman-α saturation) and one where only sources present in our simulation volume contribute to the Lyman-α background (late Lyman-α saturation). However, the most likely scenario is somewhere in between, with earlier sources, such as mini-haloes, contributing a non-negligible Lyman-α flux to the background but not achieving full Lyman-α saturation.

Despite these few caveats, the results from our simulations have shown that the 21-cm signal may not only be statistically detectable with SKA1-Low but also image-able. Our X-ray source models have distinct power spectra with markedly different evolution for our different models. Finally, the high-non Gaussianity driven by X-ray heating illustrates the need to consider statistics beyond the power spectrum when considering the 21-cm signal from the CD, particularly when considering rare X-ray sources.

Chapter 5

Conclusions

The 21-cm signal from the abundant neutral hydrogen present during the CD will provide us with a unique window into the astrophysical processes of this epoch. The upcoming radio interferometer SKA will have unprecedented sensitivity, sufficient to probe these early times. SKA will be best suited to detecting fluctuations in this signal, which is what I have focussed on in this work. Fluctuations in the signal from the CD are dominated by long-range inhomogeneous X-ray heating from HMXBs and QSOs (which were fully-numerically simulated) and variations in the Lyman-α background (which was implemented semi-numerically).

Our results, when smoothed to the resolution of SKA1-Low, show that all X-ray models considered introduce fluctuations that are not only above the rms expected noise, but also above the noise per resolution element. This means that as well as being statistically detectable, it is also plausible that the signal could be imaged.

A power spectra detection of the CD could be used to infer the presence of X-ray heating. All of our X-ray source models boost power on large scales before temperature saturation occurs. Like the power spectra, our results show that non-Gaussianity is also a good indicator of the presence of X-ray heating since all our X-ray source models have higher skewness and kurtosis values than the stellar only model where no heating occurs.

When QSOs are the only source of X-ray heating, temperature saturation occurs later than in the cases including HMXBs, although the evolution of the different cases is similar. The power spectra may be able to distinguish these two scenarios, however, it is almost completely un-affected by QSOs when they are not the dominant source of heating. On the other hand, non-Gaussianity is very high in all simulations including QSOs, even

when their heating contribution is sub-dominant. Measures of non-Gaussianity, therefore, provide a much more robust method of differentiating between X-ray sources.

However, this non-Gaussianity could be suppressed by late Lyman-α saturation. This would mean that distinguishing between X-ray models would be more difficult, although the results from our simulations show that it is still possible. Similarly, when QSOs are the only sources of X-rays, temperature saturation does not occur until reionization is underway. This can suppress the skewness and kurtosis driven by the formation of ionized bubbles, in which case it may be more difficult to detect the 21-cm signal at later times via higher order statistics.

The possibility of imaging the CD motivates the development of additional techniques for analysing the 21-cm signal from this time. Examples include deep learning of images (Gillet et al., 2018) and size distribution statistics (Giri et al., 2018). The noise is also low enough for measurements of the bispectrum to be made to further investigate non-Gaussianity in the signal (Shimabukuro et al., 2016). The bispectrum estimator presented in Watkinson et al. (2017) is currently being applied to the results from our simulations in order to further quantify the non-Gaussianity arising from our different X-ray source models.

One of the key limitations of this work is the number of X-ray source models investigated given the huge number of possibilities. Due to the uncertain parameters of these early sources, further investigation of our HMXBs and QSO models with varying luminosities and spectral indices would be merited. Our models could also be extended to be more realistic, for example, we have considered the X-ray spectra to be single power laws, whereas some observations have suggested the spectra of HMXBs and QSOs may be better described by a broken power-law (e.g. Wang, 2010). Sources considered in cosmological simulations to date have been isotropic. However, in reality this is unlikely to be the case, particularly with QSOs. We have implemented anisotropic QSO sources into C^2-Ray using Xu (2015) and are currently working on a realistic way to decide the orientation of the QSOs. Finally, there are many X-ray sources we have not yet considered, such as supernovae (Chen, 2014).

The uncertainties described above lead to a huge parameter space, which fully numerical simulations are not computationally efficient enough to explore. Our simulations can, however, be used to verify the results from faster semi-numerical methods (e.g. Santos et al., 2010; Mesinger et al., 2011; Ghara et al., 2015) via code comparison projects. These

faster methods may then be used to explore parameter space of the CD directly (e.g. Greig et al., 2016) or to train neutral networks to extract parameters (e.g. Shimabukuro and Semelin, 2017).

The contribution of minihalos to the Lyman-α background is highly uncertain. We consider the two extremes in this work: that minihalos contribute no Lyman-α photons and that minihalos saturate the Lyman-α background. However, the reality is likely to be somewhere between these two scenarios. As minihalos rely on the presence of molecular hydrogen to cool and accrete gas, hydrogen dissociating Lyman-Werner radiation is able stop them forming stars. We have run a new suite of fully numerical radiative transfer simulations including Lyman-Werner radiative transfer using the method presented in Ahn et al. (2009) in order to obtain a realistic minihalo population and hence Lyman-α background.

There is some debate over the impact of heating from the scattering of Lyman-α photons. While many works have suggested that the impact of Lyman-α heating is negligible (e.g. Chen and Miralda-Escudé, 2004; Oklopčić and Hirata, 2013) other works have indicated that this may not be the case (e.g. Chuzhoy and Shapiro, 2007; Ciardi and Salvaterra, 2007). The heating from Lyman-α photons is highly unlikely to exceed that of X-rays (Ciardi et al., 2010) but may be important particularly in the first stages of the CD. We are currently working to estimate the maximum contribution from the heating from Lyman-α scattering for a set of reasonable assumptions.

Finally, the work presented in this thesis has not yet addressed one of the most challenging aspects of analysing the 21-cm signal, namely extracting the signal from foreground contamination. The current generation of radio interferometers have been unable to detect the 21-cm signal due to the complex foregrounds, which are often many orders of magnitude higher than the 21-cm signal itself. Foreground extraction techniques fall into two categories: foreground removal and avoidance. Foreground removal (e.g. Chapman et al., 2013; Ghosh et al., 2015; Gehlot et al., 2018) focuses on accurately modelling the foregrounds so they may be subtracted from the signal, whereas foreground avoidance (e.g. Mertens et al., 2018a,b) focusses only probing a triangular region in k-space where the 21-cm signal dominates, known as the EoR window. Due to the intricate nature of these processes there are no analytical estimates to date of the impact of the foreground on the 21-cm signal. However, the outputs from the simulations presented in Chapters 3 and 4 are currently being run through the pipeline described in Chapman et al. (2015) with the

results expected soon. These results will give us an idea if the CD will be detectable after the foregrounds have been removed from the signal.

In conclusion, images of the 21-cm signal resulting from HMXBs and QSO heating at SKA1-Low resolution show fluctuations large enough to be imaged. Results from our X-ray models also showed that the long range X-ray heating can introduce strong non-Gaussianity to the signal, especially the QSO-like sources. This high non-Gaussianity is not only a useful method for detecting X-ray heating, but also an excellent probe for rare X-ray sources during the CD. With a detection of this signal almost within our reach, this is certainly an exciting time to be involved in this field. This work, and work built upon it, will help enable us to make sense of this unique observation and the complex processes driving it.

Bibliography

Agarwal, B., Khochfar, S., Johnson, J. L., Neistein, E., Dalla Vecchia, C., and Livio, M. (2012). Ubiquitous seeding of supermassive black holes by direct collapse. *MNRAS*, 425:2854–2871.

Ahn, K., Iliev, I. T., Shapiro, P. R., Mellema, G., Koda, J., and Mao, Y. (2012). Detecting the Rise and Fall of the First Stars by Their Impact on Cosmic Reionization. *ApJL*, 756:L16.

Ahn, K., Shapiro, P. R., Iliev, I. T., Mellema, G., and Pen, U.-L. (2009). The Inhomogeneous Background Of H2-Dissociating Radiation During Cosmic Reionization. *ApJ*, 695:1430–1445.

Ahn, K., Xu, H., Norman, M. L., Alvarez, M. A., and Wise, J. H. (2015). Spatially Extended 21 cm Signal from Strongly Clustered Uv and X-Ray Sources in the Early Universe. *ApJ*, 802:8.

Aldrovandi, S. M. V. and Pequignot, D. (1973). Radiative and Dielectronic Recombination Coefficients for Complex Ions. *A&A*, 25:137.

Ardaneh, K., Luo, Y., Shlosman, I., Nagamine, K., Wise, J. H., and Begelman, M. C. (2018). Direct Collapse to Supermassive Black Hole Seeds with Radiation Transfer: Cosmological Halos. *MNRAS*.

Baek, S., Semelin, B., Di Matteo, P., Revaz, Y., and Combes, F. (2010). Reionization by UV or X-ray sources. *A&A*, 523:A4.

Barkana, R. (2018). Possible interaction between baryons and dark-matter particles revealed by the first stars. *Nature*, 555:71–74.

Basu-Zych, A. R., Lehmer, B., Fragos, T., Hornschemeier, A., Yukita, M., Zezas, A., and Ptak, A. (2016). Exploring the Overabundance of ULXs in Metal- and Dust-poor Local Lyman Break Analogs. *ApJ*, 818:140.

Basu-Zych, A. R., Lehmer, B. D., Hornschemeier, A. E., Gonçalves, T. S., Fragos, T., Heckman, T. M., Overzier, R. A., Ptak, A. F., and Schiminovich, D. (2013). Evidence for Elevated X-Ray Emission in Local Lyman Break Galaxy Analogs. *ApJ*, 774:152.

Becerra, F., Marinacci, F., Bromm, V., and Hernquist, L. E. (2018). Assembly of supermassive black hole seeds. *ArXiv e-prints*.

Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., Bahcall, N. A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P. C., Doi, M., Heckman, T. M., Hennessy, G. S., Ivezić, Ž., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Nash, T., Nichol, R., Pier, J. R., Richards, G. T., Schneider, D. P., Stoughton, C., Szalay, A. S., Thakar, A. R., and York, D. G. (2001). Evidence for Reionization at z˜6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar. *Astrophysics Journal*, 122:2850–2857.

Begelman, M. C., Blandford, R. D., and Rees, M. J. (1984). Theory of extragalactic radio sources. *Reviews of Modern Physics*, 56:255–351.

Begelman, M. C., Volonteri, M., and Rees, M. J. (2006). Formation of supermassive black holes by direct collapse in pre-galactic haloes. *MNRAS*, 370:289–298.

Belikov, A. V. and Hooper, D. (2009). How dark matter reionized the Universe. *Physical Review D*, 80(3):035007.

Bolton, J. S. and Haehnelt, M. G. (2007). The observed ionization rate of the intergalactic medium and the ionizing emissivity at z¿= 5: evidence for a photon-starved and extended epoch of reionization. *MNRAS*, 382:325–341.

Borisova, E., Lilly, S. J., Cantalupo, S., Prochaska, J. X., Rakic, O., and Worseck, G. (2016). Constraining the Lifetime and Opening Angle of Quasars using Fluorescent Lyman α Emission: The Case of Q0420-388. *ApJ*, 830:120.

Bouwens, R. J., Aravena, M., Decarli, R., Walter, F., da Cunha, E., Labbé, I., Bauer, F. E., Bertoldi, F., Carilli, C., Chapman, S., Daddi, E., Hodge, J., Ivison, R. J., Karim, A., Le Fevre, O., Magnelli, B., Ota, K., Riechers, D., Smail, I. R., van der Werf, P., Weiss, A., Cox, P., Elbaz, D., Gonzalez-Lopez, J., Infante, L., Oesch, P., Wagg, J., and Wilkins, S. (2016). ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: The Infrared Excess of UV-Selected z = 2-10 Galaxies as a Function of UV-Continuum Slope and Stellar Mass. *ApJ*, 833:72.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Labbé, I., Trenti, M., van Dokkum, P., Franx, M., Stiavelli, M., Carollo, C. M., Magee, D., and Gonzalez, V. (2011). Ultraviolet Luminosity Functions from 132 z ˜ 7 and z ˜ 8 Lyman-break Galaxies in the Ultra-deep HUDF09 and Wide-area Early Release Science WFC3/IR Observations. *ApJ*, 737:90.

Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., and Mahesh, N. (2018). An absorption profile centred at 78 megahertz in the sky-averaged spectrum. *Nature*, 555:67–70.

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., and Lemson, G. (2009). Resolving cosmic structure formation with the Millennium-II Simulation. *MNRAS*, 398:1150–1164.

Brightman, M., Silverman, J. D., Mainieri, V., Ueda, Y., Schramm, M., Matsuoka, K., Nagao, T., Steinhardt, C., Kartaltepe, J., Sanders, D. B., Treister, E., Shemmer, O., Brandt, W. N., Brusa, M., Comastri, A., Ho, L. C., Lanzuisi, G., Lusso, E., Nandra, K., Salvato, M., Zamorani, G., Akiyama, M., Alexander, D. M., Bongiorno, A., Capak, P., Civano, F., Del Moro, A., Doi, A., Elvis, M., Hasinger, G., Laird, E. S., Masters, D., Mignoli, M., Ohta, K., Schawinski, K., and Taniguchi, Y. (2013). A statistical relation between the X-ray spectral index and Eddington ratio of active galactic nuclei in deep surveys. *MNRAS*, 433:2485–2496.

Bromm, V. (2013a). Formation of the first stars. *Reports on Progress in Physics*, 76(11):112901.

Bromm, V. (2013b). The first stars and galaxies - Basic principles. *Asociacion Argentina de Astronomia La Plata Argentina Book Series*, 4:3.

Bromm, V. (2017). The first stars: our evolving theoretical picture. *Mem. Societa Astronomica Italiana*, 88:671.

Bromm, V., Coppi, P. S., and Larson, R. B. (1999). Forming the First Stars in the Universe: The Fragmentation of Primordial Gas. *ApJL*, 527:L5–L8.

Bunker, A., Stanway, E., Ellis, R., Lacy, M., McMahon, R., Stark, D., Chiu, K., and Eyles, L. (2007). Galaxies at high redshift and reionization. *Nuovo Cimento B Serie*, 122:993–999.

Chandar, R., Rangelov, B., and Prestwich, A. (2009). High Mass X-Ray Binaries in Nearby Starburst Galaxies. In *American Astronomical Society Meeting Abstracts #213*,

volume 41 of *Bulletin of the American Astronomical Society*, page 329.

Chapman, E., Abdalla, F. B., Bobin, J., Starck, J.-L., Harker, G., Jelić, V., Labropoulos, P., Zaroubi, S., Brentjens, M. A., de Bruyn, A. G., and Koopmans, L. V. E. (2013). The scale of the problem: recovering images of reionization with Generalized Morphological Component Analysis. *MNRAS*, 429:165–176.

Chapman, E., Bonaldi, A., Harker, G., Jelic, V., Abdalla, F. B., Bernardi, G., Bobin, J., Dulwich, F., Mort, B., Santos, M., and Starck, J. L. (2015). Cosmic Dawn and Epoch of Reionization Foreground Removal with the SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 5.

Chen, K.-J. (2014). Supernovae at the cosmic dawn. *International Journal of Modern Physics D*, 23:1430008.

Chen, X. and Miralda-Escudé, J. (2004). The Spin-Kinetic Temperature Coupling and the Heating Rate due to Lyα Scattering before Reionization: Predictions for 21 Centimeter Emission and Absorption. *ApJ*, 602:1–11.

Chuzhoy, L. and Shapiro, P. R. (2006). Ultraviolet Pumping of Hyperfine Transitions in the Light Elements, with Application to 21 cm Hydrogen and 92 cm Deuterium Lines from the Early Universe. *ApJ*, 651:1–7.

Chuzhoy, L. and Shapiro, P. R. (2007). Heating and Cooling of the Early Intergalactic Medium by Resonance Photons. *ApJ*, 655:843–846.

Ciardi, B. and Madau, P. (2003). Probing beyond the Epoch of Hydrogen Reionization with 21 Centimeter Radiation. *Astrophysics Journal*, 596:1–8.

Ciardi, B. and Salvaterra, R. (2007). Lyα heating and its impact on early structure formation. *MNRAS*, 381:1137–1142.

Ciardi, B., Salvaterra, R., and Di Matteo, T. (2010). Lyα versus X-ray heating in the high-z intergalactic medium. *MNRAS*, 401:2635–2640.

Coles, P. and Frenk, C. S. (1991). Skewness and large-scale structure. *MNRAS*, 253:727–737.

D'Amico, G., Panci, P., Lupi, A., Bovino, S., and Silk, J. (2018). Massive black holes from dissipative dark matter. *MNRAS*, 473:328–335.

Datta, K. K., Ghara, R., Majumdar, S., Choudhury, T. R., Bharadwaj, S., Roy, H., and Datta, A. (2016). Probing Individual Sources during Reionization and Cosmic Dawn using Square Kilometre Array HI 21-cm Observations. *Journal of Astrophysics and Astronomy*, 37:27.

Davis, L. E., Cawson, M., Davies, R. L., and Illingworth, G. (1985). CCD surface photometry of galaxies with dynamical data. I - NGC 3379, M87, and NGC 1052. *AJ*, 90:169–182.

Dawoodbhoy, T., Shapiro, P. R., Choi, J.-H., Ocvirk, P., Gillet, N., Aubert, D., Iliev, I. T., Teyssier, R., Yepes, G., Sullivan, D., Knebe, A., Gottloeber, S., D'Aloisio, A., Park, H., Hoffman, Y., and Stranex, T. (2017). The Suppression of Star Formation in Low-Mass Galaxies Caused by the Reionization of their Local Patch. In *American Astronomical Society Meeting Abstracts #230*, volume 230 of *American Astronomical Society Meeting Abstracts*, page 314.02.

DeBoer, D. R., Parsons, A. R., Aguirre, J. E., Alexander, P., Ali, Z. S., Beardsley, A. P., Bernardi, G., Bowman, J. D., Bradley, R. F., Carilli, C. L., et al. (2017). Hydrogen epoch of reionization array (hera). *Publications of the Astronomical Society of the Pacific*, 129(974):045001.

Dixon, K. L., Iliev, I. T., Mellema, G., Ahn, K., and Shapiro, P. R. (2016a). The large-scale observational signatures of low-mass galaxies during reionization. *MNRAS*, 456:3011–3029.

Dixon, K. L., Iliev, I. T., Mellema, G., Ahn, K., and Shapiro, P. R. (2016b). The large-scale observational signatures of low-mass galaxies during reionization. *MNRAS*, 456:3011–3029.

Eide, M. B., Graziani, L., Ciardi, B., Feng, Y., Kakiichi, K., and Di Matteo, T. (2018). The epoch of cosmic heating by early sources of X-rays. *MNRAS*, 476:1174–1190.

Ewall-Wice, A., Chang, T.-C., Lazio, J., Doré, O., Seiffert, M., and Monsalve, R. A. (2018). Modeling the Radio Background from the First Black Holes at Cosmic Dawn: Implications for the 21 cm Absorption Amplitude. *ArXiv e-prints*.

Fan, X., Strauss, M. A., Richards, G. T., Hennawi, J. F., Becker, . H., White, R. L., Diamond-Stanic, A. M., Donley, J. L., Jiang, L., Kim, J. S., Vestergaard, M., Young, J. E., Gunn, J. E., Lupton, R. H., Knapp, G. R., Schneider, D. P., Brandt, W. N., Bahcall, N. A., Barentine, J. C., Brinkmann, J., Brewington, H. J., Fukugita, M.,

Harvanek, M., Kleinman, S. J., Krzesinski, J., Long, D., Neilsen, Jr., E. H., Nitta, A., Snedden, S. A., and Voges, W. (2006). A Survey of z¿5.7 Quasars in the Sloan Digital Sky Survey. IV. Discovery of Seven Additional Quasars. *Astrophysics Journal*, 131:1203–1209.

Fan, X., Strauss, M. A., Schneider, D. P., Becker, R. H., White, R. L., Haiman, Z., Gregg, M., Pentericci, L., Grebel, E. K., Narayanan, V. K., Loh, Y.-S., Richards, G. T., Gunn, J. E., Lupton, R. H., Knapp, G. R., Ivezić, Ž., Brandt, W. N., Collinge, M., Hao, L., Harbeck, D., Prada, F., Schaye, J., Strateva, I., Zakamska, N., Anderson, S., Brinkmann, J., Bahcall, N. A., Lamb, D. Q., Okamura, S., Szalay, A., and York, D. G. (2003). A Survey of z¿5.7 Quasars in the Sloan Digital Sky Survey. II. Discovery of Three Additional Quasars at z¿6. *Astrophysics Journal*, 125:1649–1659.

Ferrarese, L. and Merritt, D. (2000). A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. *ApJL*, 539:L9–L12.

Fialkov, A. and Barkana, R. (2014). The rich complexity of 21-cm fluctuations produced by the first stars. *MNRAS*, 445:213–224.

Fialkov, A., Barkana, R., and Visbal, E. (2014). The observable signature of late heating of the Universe during cosmic reionization. *Nature*, 506:197–199.

Field, G. B. (1958). Excitation of the Hydrogen 21-CM Line. *Proceedings of the IRE*, 46:240–250.

Field, G. B. (1959). The Spin Temperature of Intergalactic Neutral Hydrogen. *Astrophysics Journal*, 129:536.

Fioc, M. and Rocca-Volmerange, B. (1999). PEGASE.2, a metallicity-consistent spectral evolution model of galaxies: the documentation and the code. *ArXiv Astrophysics e-prints*.

Fragos, T., Lehmer, B., Tremmel, M., Tzanavaris, P., Basu-Zych, A., Belczynski, K., Hornschemeier, A., Jenkins, L., Kalogera, V., Ptak, A., and Zezas, A. (2013a). X-Ray Binary Evolution Across Cosmic Time. *ApJ*, 764:41.

Fragos, T., Lehmer, B. D., Naoz, S., Zezas, A., and Basu-Zych, A. (2013b). Energy Feedback from X-Ray Binaries in the Early Universe. *ApJL*, 776:L31.

Friedrich, M. M., Mellema, G., Iliev, I. T., and Shapiro, P. R. (2012). Radiative transfer

of energetic photons: X-rays and helium ionization in C2-RAY. *Monthly Notices of the Royal Astronomy Society*, 421:2232–2250.

Furlanetto, S. R., Oh, S. P., and Briggs, F. H. (2006). Cosmology at low frequencies: The 21 cm transition and the high-redshift universe. *Physics Reports*, 433(4-6):181–301.

Furlanetto, S. R., Oh, S. P., and Briggs, F. H. (2006). Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Phys. Reports*, 433:181–301.

Gamow, G. and Teller, E. (1939). The Expanding Universe and the Origin of the Great Nebulæ. *Nature*, 143:116–117.

Gehlot, B. K., Koopmans, L. V. E., de Bruyn, A. G., Zaroubi, S., Brentjens, M. A., Asad, K. M. B., Hatef, M., Jelić, V., Mevius, M., Offringa, A. R., Pandey, V. N., and Yatawatta, S. (2018). Wide-field LOFAR-LBA power-spectra analyses: impact of calibration, polarization leakage, and ionosphere. *MNRAS*, 478:1484–1501.

Ghara, R., Choudhury, T. R., and Datta, K. K. (2015). 21 cm signal from cosmic dawn: imprints of spin temperature fluctuations and peculiar velocities. *MNRAS*, 447:1806–1825.

Ghara, R., Choudhury, T. R., and Datta, K. K. (2016). 21-cm signature of the first sources in the Universe: prospects of detection with SKA. *MNRAS*, 460:827–843.

Ghosh, A., Koopmans, L. V. E., Chapman, E., and Jelić, V. (2015). A Bayesian analysis of redshifted 21-cm H I signal and foregrounds: simulations for LOFAR. *MNRAS*, 452:1587–1600.

Giallongo, E., Grazian, A., Fiore, F., Fontana, A., Pentericci, L., Vanzella, E., Dickinson, M., Kocevski, D., Castellano, M., Cristiani, S., Ferguson, H., Finkelstein, S., Grogin, N., Hathi, N., Koekemoer, A. M., Newman, J. A., and Salvato, M. (2015). Faint AGNs at z ¿ 4 in the CANDELS GOODS-S field: looking for contributors to the reionization of the Universe. *A&A*, 578:A83.

Gillet, N., Mesinger, A., Greig, B., Liu, A., and Ucci, G. (2018). Deep learning from 21-cm images of the Cosmic Dawn. *ArXiv e-prints*.

Giri, S. K., Mellema, G., Dixon, K. L., and Iliev, I. T. (2018). Bubble size statistics during reionization from 21-cm tomography. *MNRAS*, 473:2949–2964.

Glover, S. C. O. and Brand, P. W. J. L. (2003). Radiative feedback from an early X-ray background. *MNRAS*, 340:210–226.

Greig, B. and Mesinger, A. (2017). The global history of reionization. *MNRAS*, 465:4838–4852.

Greig, B., Mesinger, A., and Pober, J. C. (2016). Constraints on the temperature of the intergalactic medium at z = 8.4 with 21-cm observations. *MNRAS*, 455:4295–4300.

Grimm, H.-J., Gilfanov, M., and Sunyaev, R. (2003). High-mass X-ray binaries as a star formation rate indicator in distant galaxies. *MNRAS*, 339:793–809.

Gunn, J. E. and Peterson, B. A. (1965). On the Density of Neutral Hydrogen in Intergalactic Space. *Astrophysics Journal*, 142:1633–1641.

Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23:347–356.

Haiman, Z., Abel, T., and Rees, M. J. (2000). The Radiative Feedback of the First Cosmological Objects. *ApJ*, 534:11–24.

Haiman, Z. and Holder, G. P. (2003). The Reionization History at High Redshifts. I. Physical Models and New Constraints from Cosmic Microwave Background Polarization. *ApJ*, 595:1–12.

Halpern, J. P. (1984). Variable X-ray absorption in the QSO MR 2251 - 178. *ApJ*, 281:90–94.

Harnois-Déraps, J., Pen, U.-L., Iliev, I. T., Merz, H., Emberson, J. D., and Desjacques, V. (2013). High-performance P^3M N-body code: CUBEP^{3}M. *MNRAS*, 436:540–559.

Hickox, R. C. and Markevitch, M. (2007). Resolving the Unresolved Cosmic X-Ray Background in the Chandra Deep Fields. *ApJL*, 661:L117–L121.

Higgins, J. and Meiksin, A. (2012). The scattering of Lyα radiation in the intergalactic medium: numerical methods and solutions. *MNRAS*, 426:2380–2403.

Hill, J. C. and Baxter, E. J. (2018). Can Early Dark Energy Explain EDGES? *ArXiv e-prints*.

Hills, R., Kulkarni, G., Meerburg, P. D., and Puchwein, E. (2018). Concerns about Modelling of Foregrounds and the 21-cm Signal in EDGES data. *ArXiv e-prints*.

Hubble, E. and Humason, M. L. (1931). The Velocity-Distance Relation among Extra-Galactic Nebulae. *ApJ*, 74:43.

Iliev, I. T., Mellema, G., Ahn, K., Shapiro, P. R., Mao, Y., and Pen, U.-L. (2014). Simulating cosmic reionization: how large a volume is large enough? *MNRAS*, 439:725–743.

Iliev, I. T., Mellema, G., Pen, U.-L., Merz, H., Shapiro, P. R., and Alvarez, M. A. (2006). Simulating cosmic reionization at large scales - I. The geometry of reionization. *MNRAS*, 369:1625–1638.

Iliev, I. T., Mellema, G., Shapiro, P. R., and Pen, U.-L. (2007a). Self-regulated reionization. *MNRAS*, 376:534–548.

Iliev, I. T., Mellema, G., Shapiro, P. R., and Pen, U.-L. (2007b). Self-regulated reionization. *MNRAS*, 376:534–548.

Iliev, I. T., Mellema, G., Shapiro, P. R., Pen, U.-L., Mao, Y., Koda, J., and Ahn, K. (2012). Can 21-cm observations discriminate between high-mass and low-mass galaxies as reionization sources? *MNRAS*, 423:2222–2253.

Iliev, I. T., Shapiro, P. R., Ferrara, A., and Martel, H. (2002). On the Direct Detectability of the Cosmic Dark Ages: 21 Centimeter Emission from Minihalos. *ApJL*, 572:L123–L126.

Inayoshi, K., Haiman, Z., and Ostriker, J. P. (2016). Hyper-Eddington accretion flows on to massive black holes. *MNRAS*, 459:3738–3755.

Jin, C., Done, C., and Ward, M. (2017). Super-Eddington QSO RX J0439.6-5311 - I. Origin of the soft X-ray excess and structure of the inner accretion flow. *MNRAS*, 468:3663–3681.

Justham, S. and Schawinski, K. (2012). Another thread in the tapestry of stellar feedback: X-ray binaries. *MNRAS*, 423:1641–1651.

Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., and Giveon, U. (2000). Reverberation Measurements for 17 Quasars and the Size-Mass-Luminosity Relations in Active Galactic Nuclei. *ApJ*, 533:631–649.

Kawasaki, M. and Tada, Y. (2016). Can massive primordial black holes be produced in mild waterfall hybrid inflation? *Journal of Cosmology and Astroparticle Physics*, 8:041.

Khrykin, I. S., Hennawi, J. F., and McQuinn, M. (2017). The Thermal Proximity Effect: A New Probe of the He II Reionization History and Quasar Lifetime. *ApJ*, 838:96.

Kooistra, R., Silva, M. B., and Zaroubi, S. (2017). Filament hunting: integrated H i 21 cm emission from filaments inferred by galaxy surveys. *Monthly Notices of the Royal Astronomy Society*, 468:857–869.

Koopmans, L., Pritchard, J., Mellema, G., Aguirre, J., Ahn, K., Barkana, R., van Bemmel, I., Bernardi, G., Bonaldi, A., Briggs, F., de Bruyn, A. G., Chang, T. C., Chapman, E., Chen, X., Ciardi, B., Dayal, P., Ferrara, A., Fialkov, A., Fiore, F., Ichiki, K., Illiev, I. T., Inoue, S., Jelic, V., Jones, M., Lazio, J., Maio, U., Majumdar, S., Mack, K. J., Mesinger, A., Morales, M. F., Parsons, A., Pen, U. L., Santos, M., Schneider, R., Semelin, B., de Souza, R. S., Subrahmanyan, R., Takeuchi, T., Vedantham, H., Wagg, J., Webster, R., Wyithe, S., Datta, K. K., and Trott, C. (2015). The Cosmic Dawn and Epoch of Reionisation with SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 1.

Kubota, K., Yoshiura, S., Shimabukuro, H., and Takahashi, K. (2016). Expected constraints on models of the epoch of reionization with the variance and skewness in redshifted 21 cm-line fluctuations. *Publications of the Astronomical Society of Japan*, 68:61.

Lee, K.-Y., Mellema, G., and Lundqvist, P. (2016). Efficient photoheating algorithms in time-dependent photoionization simulations. *MNRAS*, 455:4406–4425.

Leitherer, C., Schaerer, D., Goldader, J. D., Delgado, R. M. G., Robert, C., Kune, D. F., de Mello, D. F., Devost, D., and Heckman, T. M. (1999). Starburst99: Synthesis Models for Galaxies with Active Star Formation. *ApJS*, 123:3–40.

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., and Peng, Y. (2013). Gas Regulation of Galaxies: The Evolution of the Cosmic Specific Star Formation Rate, the Metallicity-Mass-Star-formation Rate Relation, and the Stellar Content of Halos. *ApJ*, 772:119.

Linden, T., Kalogera, V., Sepinsky, J. F., Prestwich, A., Zezas, A., and Gallagher, J. S. (2010). The Effect of Starburst Metallicity on Bright X-ray Binary Formation Pathways. *ApJ*, 725:1984–1994.

Liu, H., Slatyer, T. R., and Zavala, J. (2016). Contributions to cosmic reionization from dark matter annihilation and decay. *Physical Review D*, 94(6):063507.

Lutovinov, A., Revnivtsev, M., Gilfanov, M., Shtykovskiy, P., Molkov, S., and Sunyaev, R. (2005). INTEGRAL insight into the inner parts of the Galaxy. High mass X-ray binaries. *A&A*, 444:821–829.

Madau, P. and Dickinson, M. (2014). Cosmic Star-Formation History. *Ann. Rev. Astron.& Astrophys.* , 52:415–486.

Madau, P. and Fragos, T. (2017). Radiation Backgrounds at Cosmic Dawn: X-Rays from Compact Binaries. *ApJ*, 840:39.

Madau, P. and Haardt, F. (2015). Cosmic Reionization after Planck: Could Quasars Do It All? *ApJL*, 813:L8.

Madau, P., Meiksin, A., and Rees, M. J. (1997). 21 Centimeter Tomography of the Intergalactic Medium at High Redshift. *ApJ*, 475:429–444.

Madau, P. and Rees, M. J. (2001). Massive Black Holes as Population III Remnants. *ApJL*, 551:L27–L30.

Marchi, F., Pentericci, L., Guaita, L., Ribeiro, B., Castellano, M., Schaerer, D., Hathi, N. P., Lemaux, B. C., Grazian, A., Le Fèvre, O., Garilli, B., Maccagni, D., Amorin, R., Bardelli, S., Cassata, P., Fontana, A., Koekemoer, A. M., Le Brun, V., Tasca, L. A. M., Thomas, R., Vanzella, E., Zamorani, G., and Zucca, E. (2017). New constraints on the average escape fraction of Lyman continuum radiation in z 4 galaxies from the VIMOS Ultra Deep Survey (VUDS). *A&A*, 601:A73.

Matthews, T. A. and Sandage, A. R. (1963). Optical Identification of 3C 48, 3C 196, and 3C 286 with Stellar Objects. *ApJ*, 138:30.

Meiksin, A. (2005). Constraints on the ionization sources of the high-redshift intergalactic medium. *MNRAS*, 356:596–606.

Meiksin, A. (2006). Energy transfer by the scattering of resonant photons. *MNRAS*, 370:2025–2037.

Meiksin, A. and Madau, P. (1993). On the photoionization of the intergalactic medium by quasars at high redshift. *ApJ*, 412:34–55.

Mellema, G., Iliev, I. T., Alvarez, M. A., and Shapiro, P. R. (2006a). C2-ray: A new method for photon-conserving transport of ionizing radiation. *Elsevier Science*, 11:374–395.

Mellema, G., Iliev, I. T., Pen, U.-L., and Shapiro, P. R. (2006b). Simulating cosmic reionization at large scales - II. The 21-cm emission features and statistical signals. *MNRAS*, 372:679–692.

Mellema, G., Koopmans, L. V., Abdalla, F. A., Bernardi, G., Ciardi, B., Daiboo, S., de Bruyn, A., Datta, K. K., Falcke, H., Ferrara, A., et al. (2013). Reionization and the cosmic dawn with the square kilometre array. *Experimental Astronomy*, 36(1-2):235–318.

Mertens, F., Ghosh, A., and Koopmans, L. V. E. (2018a). Robust Foregrounds Removal for 21-cm Experiments. In Jelić, V. and van der Hulst, T., editors, *Peering towards Cosmic Dawn*, volume 333 of *IAU Symposium*, pages 284–287.

Mertens, F. G., Ghosh, A., and Koopmans, L. V. E. (2018b). Statistical 21-cm signal separation via Gaussian Process Regression analysis. *MNRAS*, 478:3640–3652.

Mesinger, A., Ferrara, A., and Spiegel, D. S. (2013). Signatures of X-rays in the early Universe. *MNRAS*, 431:621–637.

Mesinger, A., Furlanetto, S., and Cen, R. (2011). 21CMFAST: a fast, seminumerical simulation of the high-redshift 21-cm signal. *MNRAS*, 411:955–972.

Middei, R., Vagnetti, F., Bianchi, S., La Franca, F., Paolillo, M., and Ursini, F. (2017). A long-term study of AGN X-ray variability . Structure function analysis on a ROSAT-XMM quasar sample. *A&A*, 599:A82.

Mineo, S., Gilfanov, M., Lehmer, B. D., Morrison, G. E., and Sunyaev, R. (2014). X-ray emission from star-forming galaxies - III. Calibration of the L_X-SFR relation up to redshift z $\approx$ 1.3. *Monthly Notices of the Royal Astronomical Society*, 437:1698–1707.

Mineo, S., Gilfanov, M., and Sunyaev, R. (2012a). X-ray emission from star-forming galaxies - I. High-mass X-ray binaries. *MNRAS*, 419:2095–2115.

Mineo, S., Gilfanov, M., and Sunyaev, R. (2012b). X-ray emission from star-forming galaxies - II. Hot interstellarmedium. *MNRAS*, 426:1870–1883.

Mirabel, I. F., Dijkstra, M., Laurent, P., Loeb, A., and Pritchard, J. R. (2011). Stellar black holes at the dawn of the universe. *A&A*, 528:A149.

Miralda-Escudé, J. (1993). On the He II Gunn-Peterson effect and the He II forest. *MNRAS*, 262:273–276.

Mortlock, D. J., Warren, S. J., Venemans, B. P., Patel, M., Hewett, P. C., McMahon, R. G., Simpson, C., Theuns, T., Gonzáles-Solares, E. A., Adamson, A., Dye, S., Hambly, N. C., Hirst, P., Irwin, M. J., Kuiper, E., Lawrence, A., and Röttgering, H. J. A. (2011). A luminous quasar at a redshift of z = 7.085. *Nature*, 474:616–619.

Muñoz, J. B., Dvorkin, C., and Loeb, A. (2018). 21-cm Fluctuations from Charged Dark Matter. *ArXiv e-prints*.

Oklopčić, A. and Hirata, C. M. (2013). Lyα Heating of Inhomogeneous High-redshift Intergalactic Medium. *ApJ*, 779:146.

Osterbrock, D. E. and Ferland, G. J. (2006). *Astrophysics of gaseous nebulae and active galactic nuclei*.

Pacucci, F., Mesinger, A., Mineo, S., and Ferrara, A. (2014a). The X-ray spectra of the first galaxies: 21 cm signatures. *MNRAS*, 443:678–686.

Pacucci, F., Mesinger, A., Mineo, S., and Ferrara, A. (2014b). The X-ray spectra of the first galaxies: 21 cm signatures. *MNRAS*, 443:678–686.

Pan, H. C., Stewart, G. C., and Pounds, K. A. (1990). The variable X-ray absorption and soft X-ray excess of the QSO MR2251 - 178. *MNRAS*, 242:177–187.

Pan, T., Kasen, D., and Loeb, A. (2012). Pair-instability supernovae at the epoch of reionization. *MNRAS*, 422:2701–2711.

Patil, A. H., Yatawatta, S., Koopmans, L. V. E., de Bruyn, A. G., Brentjens, M. A., Zaroubi, S., Asad, K. M. B., Hatef, M., Jelić, V., Mevius, M., Offringa, A. R., Pandey, V. N., Vedantham, H., Abdalla, F. B., Brouw, W. N., Chapman, E., Ciardi, B., Gehlot, B. K., Ghosh, A., Harker, G., Iliev, I. T., Kakiichi, K., Majumdar, S., Mellema, G., Silva, M. B., Schaye, J., Vrbanec, D., and Wijnholds, S. J. (2017). Upper Limits on the 21 cm Epoch of Reionization Power Spectrum from One Night with LOFAR. *ApJ*, 838:65.

Peebles, P. J. E. (1970). Structure of the Coma Cluster of Galaxies. *AJ*, 75:13.

Peebles, P. J. E. (1973). Comment on the Origin of Galactic Rotation. *Publications of the Astronomical Society of Japan*, 25:291.

Peebles, P. J. E. (1974). The Gravitational-Instability Picture and the Nature of the Distribution of Galaxies. *ApJL*, 189:L51.

Peebles, P. J. E. (1980). *The large-scale structure of the universe*.

Peebles, P. J. E. (1982). Primeval adiabatic perturbations - Effect of massive neutrinos. *ApJ*, 258:415–424.

Peebles, P. J. E. and Dicke, R. H. (1968). Origin of the Globular Star Clusters. *ApJ*, 154:891.

Pentericci, L., Vanzella, E., Fontana, A., Castellano, M., Treu, T., Mesinger, A., Dijkstra, M., Grazian, A., Bradač, M., Conselice, C., Cristiani, S., Dunlop, J., Galametz, A., Giavalisco, M., Giallongo, E., Koekemoer, A., McLure, R., Maiolino, R., Paris, D., and Santini, P. (2014). New Observations of z ˜ 7 Galaxies: Evidence for a Patchy Reionization. *ApJ*, 793:113.

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., Hoeg, E., Bastian, U., Bernacca, P. L., Crézé, M., Donati, F., Grenon, M., Grewing, M., van Leeuwen, F., van der Marel, H., Mignard, F., Murray, C. A., Le Poole, R. S., Schrijver, H., Turon, C., Arenou, F., Froeschlé, M., and Petersen, C. S. (1997). The HIPPARCOS Catalogue. *A&A*, 323:L49–L52.

Pezzulli, E., Volonteri, M., Schneider, R., and Valiante, R. (2017). The sustainable growth of the first black holes. *MNRAS*, 471:589–595.

Planck Collaboration, Adam, R., Aghanim, N., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Calabrese, E., Cardoso, J.-F., Carron, J., Chiang, H. C., Colombo, L. P. L., Combet, C., Comis, B., Couchot, F., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Di Valentino, E., Dickinson, C., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Fantaye, Y., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gruppuso, A., Gudmundsson, J. E., Hansen, F. K., Helou, G., Henrot-Versillé, S., Herranz, D., Hivon, E., Huang, Z., Ilić, S., Jaffe, A. H., Jones, W. C., Keihänen, E., Keskitalo, R., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J.-M., Langer, M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., López-Caniego, M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Mangilli, A., Maris, M., Martin, P. G., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Miville-Deschênes,

M.-A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Moss, A., Naselsky, P., Natoli, P., Oxborrow, C. A., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Patrizii, L., Perdereau, O., Perotto, L., Pettorino, V., Piacentini, F., Plaszczynski, S., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Racine, B., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rossetti, M., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Sirri, G., Sunyaev, R., Suur-Uski, A.-S., Tauber, J. A., Tenti, M., Toffolatti, L., Tomasi, M., Tristram, M., Trombetti, T., Valiviita, J., Van Tent, F., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., Zacchei, A., and Zonca, A. (2016). Planck intermediate results. XLVII. Planck constraints on reionization history. *A&A*, 596:A108.

Pospelov, M., Pradler, J., Ruderman, J. T., and Urbano, A. (2018). New Physics in the Rayleigh-Jeans Tail of the CMB. *ArXiv e-prints*.

Pritchard, J. R. and Furlanetto, S. R. (2007). 21-cm fluctuations from inhomogeneous X-ray heating before reionization. *MNRAS*, 376:1680–1694.

Pritchard, J. R. and Loeb, A. (2012). 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901.

Puchwein, E., Bolton, J. S., Haehnelt, M. G., Madau, P., Becker, G. D., and Haardt, F. (2015). The photoheating of the intergalactic medium in synthesis models of the UV background. *MNRAS*, 450:4081–4097.

Raga, A. C., Mellema, G., Arthur, S. J., Binette, L., Ferruit, P., and Steffen, W. (1999). 3D Transfer of the Diffuse Ionizing Radiation in ISM Flows and the Preionization of a Herbig-Haro Working Surface. *Rev. Mex. Astron. Astrofis.*, 35:123.

Rees, M. J. (1984). Black Hole Models for Active Galactic Nuclei. *Ann. Rev. Astron.& Astrophys.* , 22:471–506.

Ricotti, M., Gnedin, N. Y., and Shull, J. M. (2002). The Fate of the First Galaxies. I. Self-consistent Cosmological Simulations with Radiative Transfer. *ApJ*, 575:33–48.

Ross, H. E., Dixon, K., Iliev, I., and Mellema, G. (2018). New simulation of QSO X-ray heating during the Cosmic Dawn. In Jelić, V. and van der Hulst, T., editors, *IAU Symposium*, volume 333 of *IAU Symposium*, pages 34–38.

Ross, H. E., Dixon, K. L., Iliev, I. T., and Mellema, G. (2017). Simulating the impact of X-ray heating during the cosmic dawn. *MNRAS*, 468:3785–3797.

Roy, I., Xu, W., Qiu, J.-M., Shu, C.-W., and Fang, L.-Z. (2009). Wouthuysen-field Coupling in the 21 cm Region Around High-redshift Sources. *Astrophysics Journal*, 703:1992–2003.

Rybicki, G. B. (2006). Improved Fokker-Planck Equation for Resonance-Line Scattering. *ApJ*, 647:709–718.

Santos, M. G., Ferramacho, L., Silva, M. B., Amblard, A., and Cooray, A. (2010). Fast large volume simulations of the 21-cm signal from the reionization and pre-reionization epochs. *MNRAS*, 406:2421–2432.

Schaerer, D. (2002). On the properties of massive Population III stars and metal-free stellar populations. *A&A*, 382:28–42.

Schneider, R., Ferrara, A., Natarajan, P., and Omukai, K. (2002). First stars, very massive black holes, and metals. *The Astrophysical Journal*, 571(1):30.

Scott, D. and Rees, M. J. (1990). The 21-cm line at high redshift: a diagnostic for the origin of large scale structure. *MNRAS*, 247:510.

Semelin, B., Combes, F., and Baek, S. (2007). Lyman-alpha radiative transfer during the epoch of reionization: contribution to 21-cm signal fluctuations. *A&A*, 474:365–374.

Shakura, N. I. and Sunyaev, R. A. (1973). Black holes in binary systems. Observational appearance. *A&A*, 24:337–355.

Shapiro, P. R., Iliev, I. T., and Raga, A. C. (2004). Photoevaporation of cosmological minihaloes during reionization. *MNRAS*, 348:753–782.

Shimabukuro, H. and Semelin, B. (2017). Analysing the 21 cm signal from the epoch of reionization with artificial neural networks. *MNRAS*, 468:3869–3877.

Shimabukuro, H., Yoshiura, S., Takahashi, K., Yokoyama, S., and Ichiki, K. (2015). Studying 21cm power spectrum with one-point statistics. *MNRAS*, 451:467–474.

Shimabukuro, H., Yoshiura, S., Takahashi, K., Yokoyama, S., and Ichiki, K. (2016). 21 cm line bispectrum as a method to probe cosmic dawn and epoch of reionization. *MNRAS*, 458:3003–3011.

Shivaei, I., Reddy, N. A., Siana, B., Shapley, A. E., Kriek, M., Mobasher, B., Freeman, W. R., Sanders, R. L., Coil, A. L., Price, S. H., Fetherolf, T., Azadi, M., Leung, G.,

and Zick, T. (2018). The MOSDEF Survey: Direct Observational Constraints on the Ionizing Photon Production Efficiency, ξ_{ion}, at z 2. *ApJ*, 855:42.

Shtykovskiy, P. E. and Gilfanov, M. R. (2007). High-mass X-ray binaries and recent star formation history of the Small Magellanic Cloud. *Astronomy Letters*, 33:437–454.

Smith, F. J. (1966). Hydrogen atom spin-change collisions. *Planetary Space Science*, 14:929–936.

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636.

Steidel, C. C., Pettini, M., and Hamilton, D. (1995). Lyman Imaging of High-Redshift Galaxies.III.New Observations of Four QSO Fields. *AJ*, 110:2519.

Sullivan, D. and Iliev, I. T. (2016). Radiative Feedback Effects during Cosmic Reionization. In van de Weygaert, R., Shandarin, S., Saar, E., and Einasto, J., editors, *The Zeldovich Universe: Genesis and Growth of the Cosmic Web*, volume 308 of *IAU Symposium*, pages 372–377.

Sun, L., Paschalidis, V., Ruiz, M., and Shapiro, S. L. (2017). Magnetorotational collapse of supermassive stars: Black hole formation, gravitational waves, and jets. *Physical Review D*, 96(4):043006.

Tozzi, P., Gilli, R., Mainieri, V., Norman, C., Risaliti, G., Rosati, P., Bergeron, J., Borgani, S., Giacconi, R., Hasinger, G., Nonino, M., Streblyanska, A., Szokoly, G., Wang, J. X., and Zheng, W. (2006). X-ray spectral properties of active galactic nuclei in the Chandra Deep Field South. *A&A*, 451:457–474.

Tozzi, P., Madau, P., Meiksin, A., and Rees, M. J. (2000). Radio Signatures of H I at High Redshift: Mapping the End of the "Dark Ages". *ApJ*, 528:597–606.

Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., and Watson, M. G. (2014). Toward the Standard Population Synthesis Model of the X-Ray Background: Evolution of X-Ray Luminosity and Absorption Functions of Active Galactic Nuclei Including Compton-thick Populations. *ApJ*, 786:104.

Vonlanthen, P., Semelin, B., Baek, S., and Revaz, Y. (2011). Distinctive rings in the 21 cm signal of the epoch of reionization. *A&A*, 532:A97.

Wang, W. (2010). INTEGRAL/IBIS observations of a hard X-ray outburst in high-mass X-ray binary 4U 2206+54. *A&A*, 520:A22.

Watkinson, C. A., Giri, S., Ross, H. E., Pritchard, J. R., Mellema, G., and Iliev, I. T. (2018). The 21cm bispectrum as a probe of non-Gaussianities due to X-ray heating.

Watkinson, C. A., Majumdar, S., Pritchard, J. R., and Mondal, R. (2017). A fast estimator for the bispectrum and beyond - a practical method for measuring non-Gaussianity in 21-cm maps. *MNRAS*, 472:2436–2446.

Watkinson, C. A. and Pritchard, J. R. (2015). The impact of spin-temperature fluctuations on the 21-cm moments. *MNRAS*, 454:1416–1431.

Whalen, D. J., Even, W., Frey, L. H., Smidt, J., Johnson, J. L., Lovekin, C. C., Fryer, C. L., Stiavelli, M., Holz, D. E., Heger, A., Woosley, S. E., and Hungerford, A. L. (2013a). Finding the First Cosmic Explosions. I. Pair-instability Supernovae. *ApJ*, 777:110.

Whalen, D. J., Joggerst, C. C., Fryer, C. L., Stiavelli, M., Heger, A., and Holz, D. E. (2013b). Finding the First Cosmic Explosions. II. Core-collapse Supernovae. *ApJ*, 768:95.

Willott, C. J., Carilli, C. L., Wagg, J., and Wang, R. (2015). Star Formation and the Interstellar Medium in z ¿ 6 UV-luminous Lyman-break Galaxies. *ApJ*, 807:180.

Woo, J.-H. and Urry, C. M. (2002). Active Galactic Nucleus Black Hole Masses and Bolometric Luminosities. *ApJ*, 579:530–544.

Wouthuysen, S. A. (1952a). On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line. *AJ*, 57:31–32.

Wouthuysen, S. A. (1952b). On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line. *Astrophysics Journal*, 57:31–32.

Xu, H., Ahn, K., Wise, J. H., Norman, M. L., and O'Shea, B. W. (2014). Heating the Intergalactic Medium by X-Rays from Population III Binaries in High-redshift Galaxies. *ApJ*, 791:110.

Xu, Y.-D. (2015). Anisotropic radiation from accretion disc coronae in active galactic nuclei. *MNRAS*, 449:191–198.

Yajima, H. and Khochfar, S. (2015). Can the 21-cm signal probe Population III and II star formation? *MNRAS*, 448:654–665.

Yajima, H. and Li, Y. (2014). Distinctive 21-cm structures of the first stars, galaxies and quasars. *MNRAS*, 445:3674–3684.

Yang, Y. (2018). The contributions of dark matter annihilation to the global 21cm spectrum observed by the EDGES experiment. *ArXiv e-prints*.

Zaroubi, S. (2013). The Epoch of Reionization. In Wiklind, T., Mobasher, B., and Bromm, V., editors, *The First Galaxies*, volume 396 of *Astrophysics and Space Science Library*, page 45.

Appendix A

HMXB Temperature Corrections

A.1 Calculation

This explanation has been taken from Ross et al. (2017), where it has been published.

H II regions can have sizes smaller than our cell resolution, particularly for individual weak sources, and therefore be unresolved in our simulations. The cells containing such ionized regions will appear partially ionized in the simulation, with a temperature that is averaged between the hot, ionized gas phase and the colder, neutral one. At 21-cm, using the average cell T would yield a signal in emission where it should appear in absorption. In order to correct for such un-physical behaviour, we have adopted an algorithm to locate such cells and calculate $\delta T_{\rm b}$ appropriately, as follows. Since this does not affect any of the physical quantities produced by the code, it can be performed as a post-processing step.

Finding and marking the cells requiring special treatment: Since the softer stellar spectra do not produce photons that can penetrate into the IGM (the typical mean free paths are of order kpc), the cells potentially requiring correction in the stellar-only simulation are identified as those with $T > T_{\rm ad}$ and $x > x_{\rm in}$, where $T_{\rm ad}$ is the mean adiabatic gas temperature of the universe and $x_{\rm in}$ is the initial ionized fraction.

In the HMXB simulation the cells that might need correction are in the same locations as in the stellar-only simulation since the ionizing sources are identical between the two simulations and the co-located HMXBs do not contribute enough ionizing photons to

significantly grow the primarily stellar radiation-driven ionized regions (which we have tested using high resolution simulations and analytic estimates).

Calculating the temperature of H II regions in the stellar-only simulation: The softer stellar spectra yield sharp H II region boundaries, separating them from the cold, adiabatically cooling IGM (due to the lack of shock heating in our simulations, the treatment of which we leave for future work). We thus assume that the temperature of the neutral gas is the adiabatic temperature, $T_{\rm HI,s} = T_{\rm ad}$. The temperature of the H II regions in each marked cell is thus calculated using:

$$T_{\rm HII,s} = \frac{T_{\rm c,s} - T_{\rm ad}(1-x)}{x}, \tag{A.1}$$

where $T_{\rm HII,s}$ is the temperature of the H II region, $T_{\rm c,s}$ is the cell-average temperature given by C^2-Ray, $T_{\rm ad}$ is the adiabatic temperature of the universe and x is the volume weighted ionized fraction of H. Note that as marked cells have stars inside them the ionized fraction will never tend to zero so this equation will not diverge.

Calculating $\rm T_{HI}$ *from HMXB simulation:* The temperature of the neutral IGM in each marked cell in the HMXBs simulation is calculated using:

$$T_{\rm HI,x} = \frac{T_{\rm c,x} - T_{\rm HII,s}x}{1-x}, \tag{A.2}$$

where $T_{\rm c,x}$ is the temperature from C^2-RAY from the HMXB simulation and $T_{\rm HII,s}$ is from Equation A.1. Here, we assume that the temperature of the H II regions is the same between the two simulations, since the local heating in the H II region is strongly dominated by the stellar emission. This similarity was verified by high-resolution tests and analytical estimates, which showed that the additional X-ray heating is negligible.

In some cases, the LMACHs are suppressed in the HMXB case, but not the stellar-only case due to a marginally higher ionized fraction. These very rare cells are taken to be the average temperature of their neighbours. In summary, we use the temperatures of the neutral gas in each cell ($T_{\rm HI,s}$ and $T_{\rm HI,x}$, respectively for the two simulations), as calculated above to derive the 21-cm $\delta T_{\rm b}$.

A.2 Test boxes

To demonstrate the difference between the $T_{\rm K}$ and $\delta T_{\rm b}$ we have run two small test boxes with sources typical of our simulations inside. The test box a resolution of resolution as

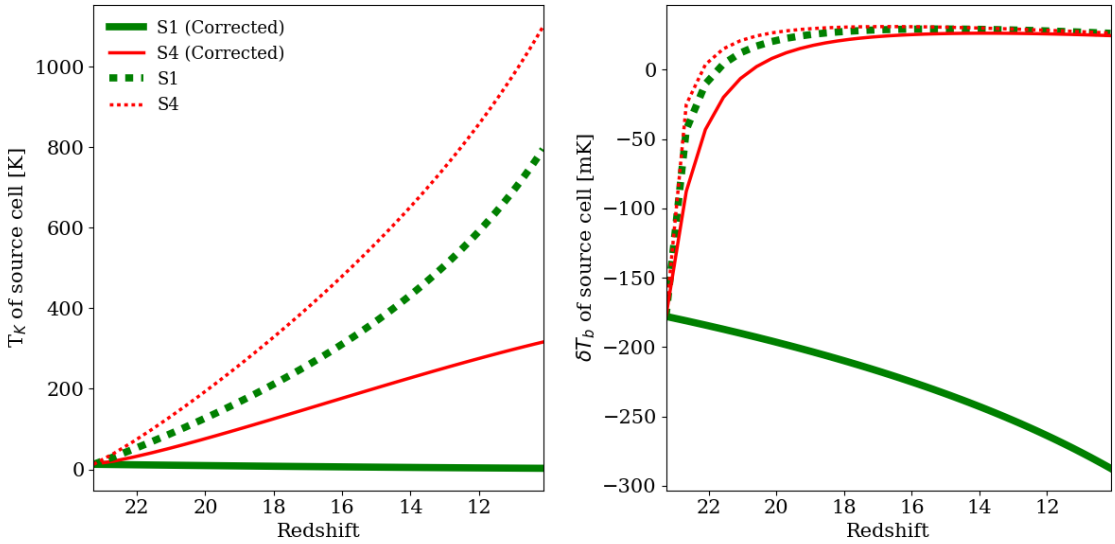

Figure A.1: Test results comparing the temperature between the new method and the old method. On the left we show the ionized fraction evolution with the 'classic' method (dashed lines) and the new multiphase method (solid lines). In the middle panel we show the mean temperature evolution using the new multiphase method (solid lines), the 'classic' method (dotted lines) and the 'classic' method with corrected temperature as done in Paper I (dashed lines). Finally, on the right we show $\delta T_{\rm b}$, following the same notation as the middle plot.

our simulations (0.976 Mpc h^{-1} per cell) and a typical density of a source cell (2 σ). (Boxes of varying density were run. The density was found to have little impact on the results.) Both boxes have a single, stellar-only source in the box centre with a luminosity typical of a stellar source in our simulations (10^{50} ionizing photons per second). The second case has a HMXB source, again with a luminosity roughly average for our simultion (10^{49} ionizing photons per second).

In the left-hand side panel of Figure A.1 $T_{\rm K}$ of the source cell is shown. Dashed lines represent $T_{\rm c,x}$ and the solid lines $T_{\rm HI,x}$ with the stellar only case in green and the HMXB case in red. Here we clearly see that the temperature given by C^2-RAY is far higher than the more accurate temperature found by our correction method. In the right-hand side panel of Figure A.1 the impact of this overestimation on $\delta T_{\rm b}$ can be seen. The stellar only case has a vastly different evolution with $\delta T_{\rm b}$ being found to be in strong emission rather than absorption. In the case of the HMXB the difference is smaller, but certainly not negligible. If the temperature output by C^2-RAY is taken to be the kinetic temperature, temperature saturation occurs almost instantaneously.

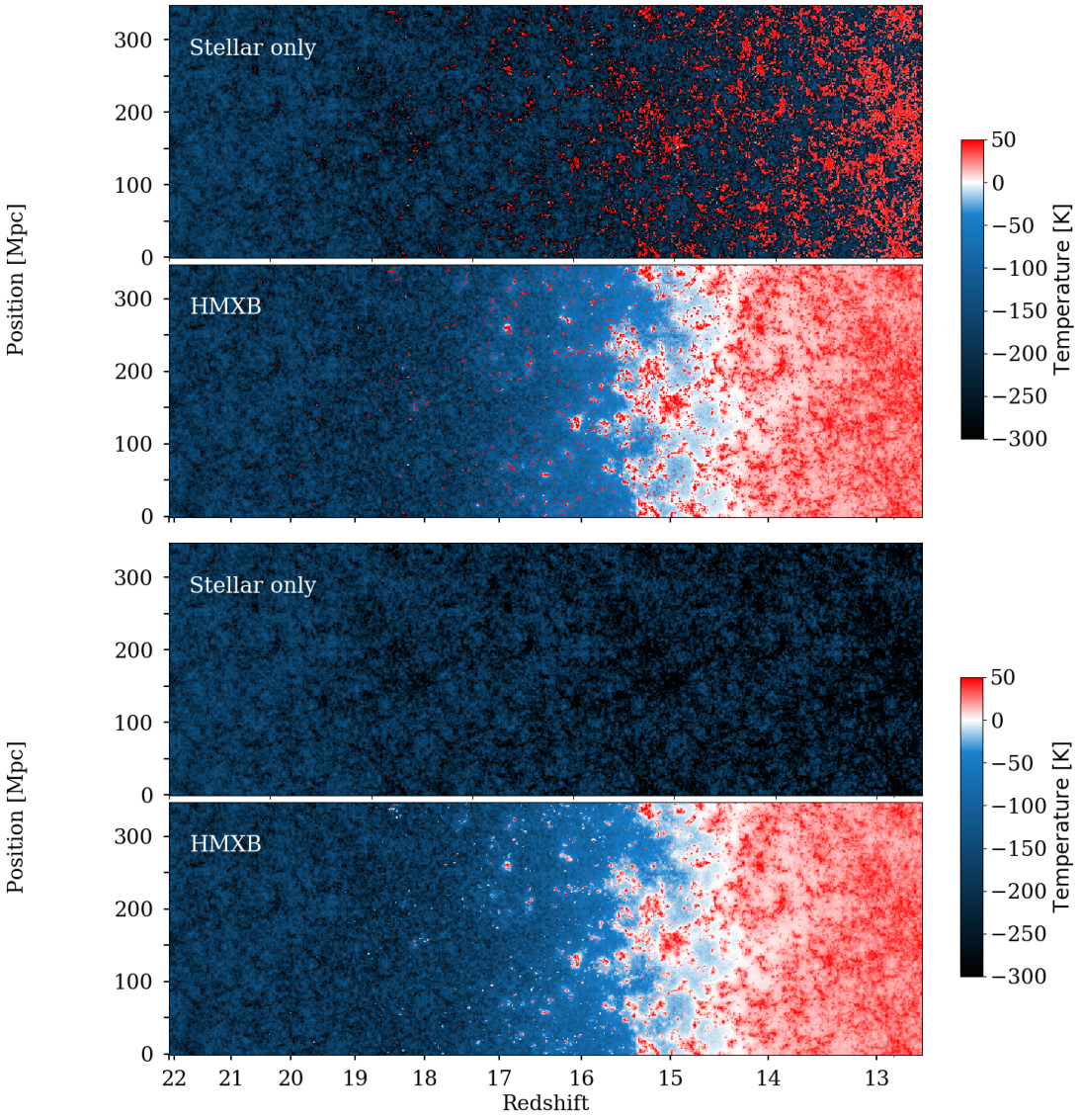

Figure A.2: Figure showing lightcones of $\delta T_{\rm b}$ before (top two panels) and after (lower two panels) the corrections are applied. As the number of source cells increases more partially ionized cells appear, the difference between the corrected and uncorrected becomes more pronounced.

A.3 Simulation results

As this effect only impacts source cells, it does not alter the signal as significantly at earlier redshifts when sources are rarer. However, these incorrect hot regions produce a signal of strong emission in the few source cells present, which adds un-physical fluctuations to the signal. By the end of the simulation volume sources exist in almost every cell, in the HMXB case this does not impact the 21-cm signal as heating saturation has occurred. However, in the stellar only case this causes an enormous inaccuracy.

Appendix B

The Addition of QSOs

Originally, C^2-Ray only contained the HMXB sources described in Section 2.2.2 and had to be updated in order to include QSO sources with different source properties. There are two key differences between HMXBs and QSO-like sources:

- QSO-like sources are not active in every halo at all times, but switch on and off. The luminosity varies while QSOs are shining.

- The luminosity of QSOs is not proportional to the halo mass but decided by additional information in the source-lists.

B.1 QSO Luminosities

Previously the sources in C^2-Ray calculate the luminosity from the mass of the host halo, as described in Section 2.2.2, which is given as a column in the halo lists. QSO sources have luminosities that do not depend on the halo mass but are instead calculated from an extrapolation of a QLF (as described in Section 2.2.3). In our updated version of C^2-Ray QSO luminosities are input as an additional column of the halo lists. Different halo-lists are therefore required for different QSO models. In order to ensure the correct halo-list is being read in the power-law index of the QSO included in the halo-list file. This value is read in with the sources and checked against the value being used in C^2-Ray.

B.2 Switching X-rays on and off

Previously, the only option the version of C^2-Ray presented in Friedrich et al. (2012) had was to run with both HMXBs and stars. In addition to implementing QSOs C^2-Ray has been updated so that both QSOs and HMXBs can be switched on and off using a flag in the Makefile. Simulations can now be be run with different combinations of X-ray sources (or without them). Preprocessing statements are used to allow conditional checks to be done at compile time rather than at runtime.

B.3 Speeding up QSO-only simulations

When a simulation does not contain X-ray sources, it is not necessary to ray trace across the entire box. In the first version of C^2-Ray the ionization rates were calculated until the vast majority of the photons had been absorbed. Therefore, particularly at the beginning of reionization when the simulation box remains optically thick, race-tracing is done over a very small number of cells. However, UV photons vastly outnumber X-ray photons. If we were to stop rays once most photons had been absorbed we would prevent long range heating from taking place. For this reason, simulations including X-ray sources are much more computationally expensive than those without as every ray has to be traced across the entire box.

However, in the case of QSO-only simulations tracing the ray across the entire box is only necessary for sources including QSOs. Radiation from the remainder of the sources need only be traced until most of the photons have been absorbed as it is entirely UV radiation. In order to speed up simulations where not every source contains an X-ray, I introduced the condition that rays only need be traced across the entire box for sources containing QSOs or HMXBs. This has been thoroughly tested and is checked at runtime (reporting any errors to the logfile).

Appendix C

The Multiphase Code

C.1 Multiphase algorithm

Many mesh based calculations adopt the finite-volume approach which means that the value of a quantity Q inside a cell is the average of this quantity over the volume of the cell V

$$\langle Q \rangle = \int_V Q \mathrm{d}V \,. \tag{C.1}$$

Problems arise when derived quantities rely non-linearly on one or more calculated quantities Q_i, as generally

$$\langle Q^k \rangle \neq \langle Q \rangle^k \,, \tag{C.2}$$

if $k \neq 1$. Examples of this in the context of photo-ionization calculations are the recombinations rates and collisional cooling rates (both are proportional to n^2, where n is the density). In the context of recombination rates, this discrepancy can sometimes be corrected for by using clumping factors

$$C = \frac{\langle n^2 \rangle}{\langle n \rangle^2} \,, \tag{C.3}$$

if the density variations inside a cell are known.

The width of ionization fronts (I-fronts) is approximately 20 mean free paths, which for soft (low-energy) photons is typically much smaller than the spatial resolution in cosmological-volume simulations. Therefore, the I-front transition is quite sharp and cannot be easily resolved. When the radiative transfer code does not resolve ionization fronts, some cells will be partly inside an ionized region, where the hydrogen ionization fraction $x \approx 1$, and partly outside, where $x \approx 0$. Let us assume that a fraction f of a cell is inside and $1 - f$

outside the ionized region. Such a cell can be described as multiphase, as it contains both an ionized phase and a neutral phase. Since for pure hydrogen the recombination rate is proportional to $n(\mathrm{HII})n(e)$ and $n(\mathrm{HII}) = n(\mathrm{e}^-) = xn$, the average recombination rate in the cell will be proportional to fn^2; whereas, the value derived from the finite volume values of the cell will be f^2n^2, as $\langle n(\mathrm{HII})\rangle = fn$. This error is usually ignored as it is often transient, around the time the I-front passes through the cell, and relatively small. However, for large cells and weak sources, cells may be in a multiphase state for a long time, and the cumulative error in the recombination calculation may be substantial. These errors can be much more substantial when calculating the 21-cm signal. The averaged value of the 21-cm signal in a cell is

$$\langle \delta T_\mathrm{b}\rangle = \delta\hat{T}_\mathrm{b}\left\langle x(\mathrm{HI})(1+\delta)\left(1-\frac{T_\mathrm{CMB}}{T_\mathrm{S}}\right)\right\rangle ; \tag{C.4}$$

whereas, the finite volume values for the gridded quantities give the estimate

$$\delta T'_\mathrm{b} = \delta\hat{T}_\mathrm{b}\langle x(\mathrm{HI})\rangle\langle(1+\delta)\rangle\left(1-\frac{T_\mathrm{CMB}}{\langle T_\mathrm{S}\rangle}\right) . \tag{C.5}$$

If we assume that $T_\mathrm{S} = T$, a fraction f a cell is fully ionized and hot ($T = T^\mathrm{hot}$), and the remainder neutral and cold ($T = T^\mathrm{cold} < T_\mathrm{CMB}$), the average 21-cm signal will be

$$\langle \delta T_\mathrm{b}\rangle = \delta\hat{T}_\mathrm{b}(1-f)(1+\delta)\left(1-\frac{T_\mathrm{CMB}}{T^\mathrm{cold}}\right) < 0. \tag{C.6}$$

However, the quantity $\delta T'_\mathrm{b}$, based on the cell averages, will be

$$\delta T'_\mathrm{b} = \delta\hat{T}_\mathrm{b}(1-f)(1+\delta)\left(1-\frac{T_\mathrm{CMB}}{fT^\mathrm{hot}+(1-f)T^\mathrm{cold}}\right) \tag{C.7}$$

and positive if $fT^\mathrm{hot} + (1-f)T^\mathrm{cold} > T_\mathrm{CMB}$, which around $z \sim 15$ and for $T^\mathrm{hot} \sim 10^4$ K is true for any $f > 4\times10^{-3}$. Clearly, we need to separate the hot and cold states to obtain the correct 21-cm signal.

Such sharp transitions in the ionization fraction and temperature are associated with relatively soft (i.e., low-energy) ionizing photons for which the mean free path is shortest. If the sources can produce both soft and hard (high-energy) ionizing photons, as is the case in the simulations presented in this paper, we need to separate their effects to establish what fraction of a cell is fully ionized by soft ionizing photons and what fraction of the cell is cold and neutral, or is partially ionized by hard photons. In Paper I, we achieved this separation by running two simulations, one with only soft sources and one with both soft and hard sources included. The difference between the two results would then give us the heating and ionization caused by the hard sources alone.

However, while this worked well for the case considered in Paper I, this solution neglects the fact that the code uses the average temperatures and electron fractions to calculate recombination and cooling rates, which are also strongly non-linear functions of the temperature. In addition, it is rather wasteful to run two simulations for every case. We therefore have introduced a different approach in which we let the code internally and self-consistently take into account the multiphase character of the ionization front cells.

This new multiphase code separates and individually tracks the effects of the photoionization and heating rates from the soft and hard sources, $\Gamma^{\rm soft}$ and $\Gamma^{\rm hard}$, and also keeps track of the mass fraction f of a cell that has been ionized by soft sources. Since soft photons have very short mean free paths in neutral or partially ionized media due to their high interaction cross-sections, there is always a sharp boundary outlining the volume affected by them. Cells can be in one of three categories:

1. Pre-multiphase cells: these cells have never seen a soft ionizing rate above certain minimum value $\Gamma^{\rm soft} > \Gamma^{\rm soft,lim}$ and are considered to not be affected by soft ionizing photons. They have a uniform ionization fraction and temperature, determined by $\Gamma^{\rm hard}$. These cells produce a 21-cm signal, calculated based on the cell-averaged quantities.

2. Multiphase cells: these cells have at some point seen a soft ionization rate $\Gamma^{\rm soft} > \Gamma^{\rm soft,lim}$ and a fraction f of these cells is assumed to be fully ionized ($x({\rm HI}) = 0$) and heated to $T^{\rm soft}$. We take $T^{\rm soft} = 10^4$ K. The value of f depends on the evolution of $\Gamma^{\rm soft}$. The ionization and thermal state of the rest of the cell, $1 - f$, is calculated based on the hard-photon rates $\Gamma^{\rm hard}$. Only this $1 - f$ cell fraction produces a 21-cm signal.

3. Post-multiphase cells: these cells have become fully ionized by soft ionizing sources, or in other words $f \approx 1$. In practice we set a limit of $f^{\rm lim} = 0.999$ above which we consider a cell to have reached this state. We calculate the ionization fraction and temperature based on the total rates $\Gamma^{\rm soft} + \Gamma^{\rm hard}$. These cells do not produce any appreciable 21-cm signal.

The minimal ionization rate needed for cells to become multiphase is set to $\Gamma^{\rm soft,lim} = 10^{-5}/\Delta t$, where Δt is the time step. This threshold implies that the soft ionizing photons can ionize a fraction $f = 10^{-5}$ of the cell within one time step. We found in tests that for the resolution and time step we use (~ 1 Mpc and ~ 5 Myr), this limit correctly identifies ionization front cells.

For the multiphase cells, we only calculate the hydrogen ionization fraction, which we equate to the fraction f that has $x(\mathrm{HII}) = 1$. We assume that inside this fraction f of the cell affected by soft photons the helium ionization state $[x(\mathrm{HeI}), x(\mathrm{HeII}), x(\mathrm{HeIII})]$ will be $[0, 1 - x(\mathrm{HeIII}), x(\mathrm{HeIII})]$, where $x(\mathrm{HeIII})$ is set by the hard sources and is assumed to have this value over the entire cell.

The code first considers the effect of the soft ionizing sources. By ray tracing from these sources and calculating the ionization rates from them, we establish which cells are in which categories. For the cells that are in the multiphase category, we find the value of f. Since we assume that HeII follows HII (which is valid as stars dominate HeII and HII ionization) and that the temperature is T^{soft}, we only calculate the hydrogen ionization rates during this step. Because we have multiple soft sources contributing to a cell, we iterate to obtain converged values for f. The recombination rates are calculated assuming the temperature T^{soft}.

After this step, we ray trace again, but now considering both hard and soft sources, calculating the H and He photo-ionization rates as well as the heating rates. For the pre- and post-multiphase cells, we apply the sum of the contributions of all sources. For the multiphase cells, we disregard the contribution of the soft sources since their contribution to the ionization was accounted for in the first step.

C.2 Description of test simulation results

In order to test and verify our new method we have run a series of simplified test cases, as summarized below.

C.2.1 Test 1: Comparison to a High-Resolution box

Firstly, we compare the raw outputs of the multiphase and old version of the code both at high and simulation resolutions (3.25 kpc h^{-1} per cell and 0.976 Mpc h^{-1} per cell, respectively). Boxes have a single, stellar-only source in the box centre with the average luminosity of a stellar source in our simulations (10^{50} ionizing photons per second), and the test is run from z=23.268 until z=22.67.

In Fig. C.1, we show the result from Test 1 for $T_{\mathrm{c,x}}$ and δT_{b} (bottom) cross-sections through the forming HII region. The results from the high-resolution box (thin red line)

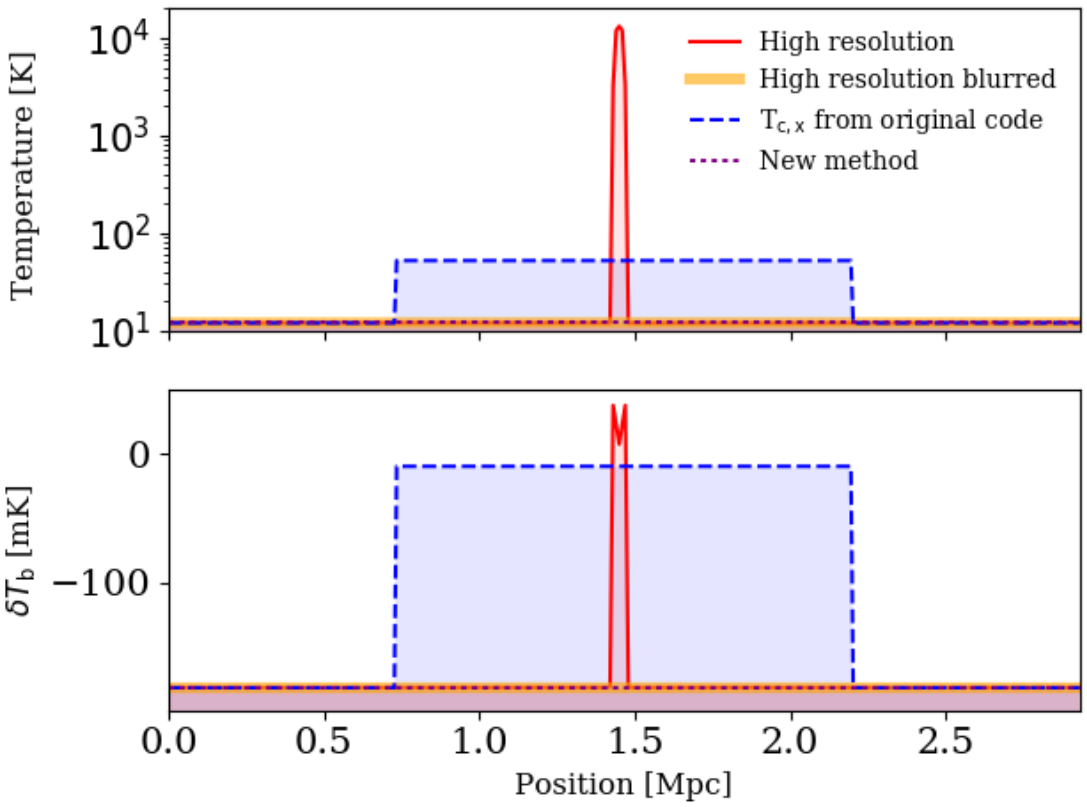

Figure C.1: Test 1: Comparison of our new multiphase radiative transfer method to the previous, 'standard' C^2-RAY method, as well as to a high resolution simulation using the old method for a typical stellar-only source at z=22.67. The plot shows cross section of the kinetic temperature (top) and δT_b (bottom). The temperature is over estimated in the original method when compared to the high resolution case. The multiphase method, however, yields the same result as the high resolution run when it is smoothed to the resolution of the simulation. In the lower panel we can see that the old method predicts a signal of emission whereas the new method and smoothed high resolution run give the expected result for a source with no X-rays i.e. absorption from the cold gas in the surrounding cells.

are correct despite using the original method since the ionization front has been fairly well resolved. Currently, computational resources are insufficient to run a full simulation box at this resolution; otherwise, high resolution runs would be a valid solution.

In order to compare the results at the resolution of our large-scale simulations, we coarsefn these results to the relevant resolution (thick orange line). The original method (dashed blue line) clearly over-estimates the temperature of the neutral IGM when compared to the smoothed high-resolution run. However, the multiphase method (dotted purple line) is in agreement with the smoothed high resolution box.

Fig. C.1 (lower panel) shows the impact this overestimation of the temperature of the neutral IGM has on $\delta T_{\rm b}$. The old code (blue dashed line) predicts $\delta T_{\rm b}$ to be much higher, and partly in emission, than the value calculated from the high resolution box then smoothed to the resolution of the simulation (thick orange line). The multiphase method (dotted purple line) again agrees with the smoothed high-resolution box (thick orange line). Note that the high resolution run does in fact show a small amount of emission due to photons from the harder end of the blackbody spectrum and (still) insufficient resolution to fully resolve the I-front. However, clearly these heated regions do not contribute significantly to $\delta T_{\rm b}$.

C.2.2 Test 2: Comparison to $\delta T_{\rm b}$ calculation used in Paper I

To compare the new multiphase method to the correction method we used in Paper I, we have performed two additional constant density test simulations at a resolution of 2σ (the effect of varying the density of the boxes was checked and concluded to be minimal). All boxes have luminous sources in the centre with the same parameters as in Table 4.1: only stars (S1), and in combination with HMXBs (S4), QSOs (S2 and S3) and all types (S5). The luminosity of HMXBs were typical of our simulations (10^{49} ionizing photons per second). QSO sources have a luminosity of 10^{52} ionizing photons per second and do not switch on until $z \sim 20$. As in the simulation, they are active for 34.5 Myr.

The results of these tests are shown in Fig. C.2 alongside results from the lower resolution boxes considered in Test 1. As the multiphase and original method treat the cells without stellar radiation equivalently, we focus on the values of the source cell only (in both of these tests the ionized bubble never leaves this cell). In Fig. C.2 (top left), we show the hydrogen ionized fraction evolution produced by the two methods. The ionized fraction of hydrogen is slightly lower in the original case for all test boxes, which is due to the fact

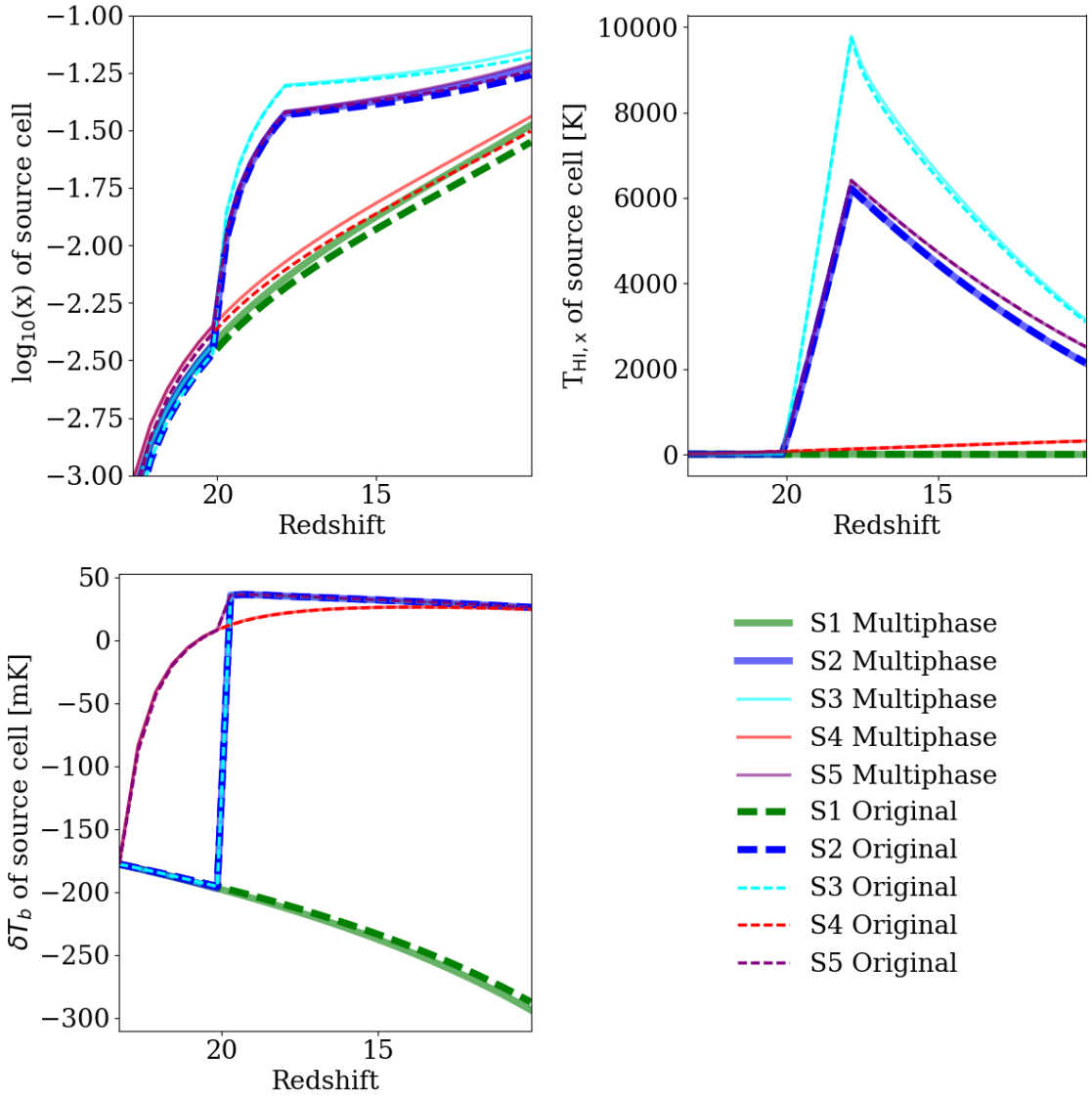

Figure C.2: Test results comparing the new multiphase method and the old method. On the top left we show the ionized fraction evolution with the 'classic' method (dashed lines) and the new multiphase method (solid lines). In the top right panel we show the mean temperature evolution using the new multiphase method (solid lines) and the 'classic' method with corrected temperature as done in Paper I (dashed lines). Finally, on bottom left we show $\delta T_{\rm b}$, following the same notation as the top right plot.

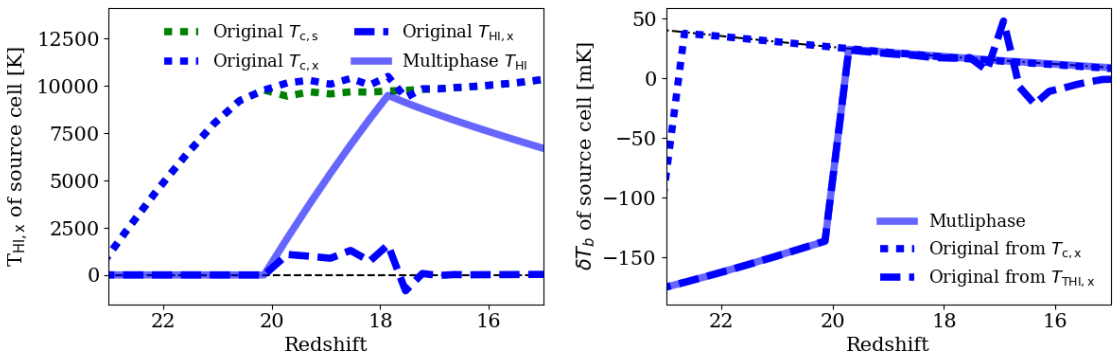

Figure C.3: Test runs showing the limitations of the old method. Here, a bright star (1×10^{52} ionizing photons per second) raises the temperature to 10,000 K. The resulting collisional cooling prevents X-ray heating from increasing the temperature and causes the temperature of the neutral IGM to be underestimated.

that recombination rates were previously calculated from the average temperature and hydrogen ionized fraction of the cell and therefore were systematically over-estimated. In the new method, recombination rates are calculated for ionized hydrogen at the assumed temperature of the ionized region (10^4 K). Apart from this minor difference, both original and multiphase methods display the same smooth evolution.

In Fig. C.2 (top right), we show the evolution of the $T_{\rm HI,x}$ of the source cell, comparing the new multiphase method (solid) and the original method (calculated as described in Paper I). We can see that, for all sources, the two methods are in agreement. In the bottom left panel of this figure, we show $\delta T_{\rm b}$ for both methods and all sources. Again, the two methods are in excellent agreement. We therefore conclude that the multiphase method and our previous method of calculating $T_{\rm K}$ using post processing can both be used in the case of HMXBs. However, it is important to note that the multiphase method is far more efficient, as it does not require an additional simulation.

C.2.3 Test 3: Limitations to the Original Method

The two methods diverge from one another when the temperature rises above 10,000 K. At these high temperatures, collisional cooling is able to lower the temperature. This cooling effectively caps the temperature, meaning that the two simulations have similar temperatures and X-ray heating is underestimated. In addition, if the X-ray simulation is slightly hotter cooling occurs more rapidly once sources are switched off. Here, $T_{\rm HII,s}$ can even become larger than $T_{\rm c,x}$, meaning that the original method yields negative tempera-

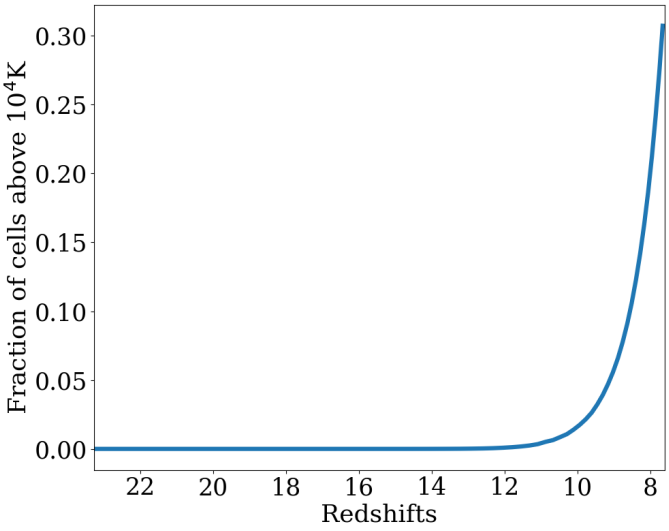

Figure C.4: Plot showing how many cells rise about 10^4K in the stellar simulation using the old version of the C^2-Ray, and so could not be corrected by the original correction method outlined in Appendix A. The number of these hot cells remains negligible until z≈12 at which point it begins to rise.

tures!

We have demonstrated this effect in a simulation-resolution test box with a bright star (1 $\times 10^{52}$ ionizing photons per second) and a QSO (the same QSO as S2 in Test 2) inside. The resulting $T_{\rm K}$ and $\delta T_{\rm b}$ are shown in Fig. C.3 in the left and right panels, respectively. The methods agree until the QSO switches on at $z \approx 20$ and the heating from the QSO is counteracted by the collisional cooling, thereby not raising $T_{\rm K}$ significantly. X-ray heating from the QSO is vastly underestimated and becomes negative after the QSO switches off, when collisional cooling is able to lower the temperature even further. The multiphase method, however, is able to smoothly evolve this cell, as it calculates the temperature self-consistently.

The fraction of cells impacted by this behaviour is plotted in Figure C.4. The number of 'hot cells' in redshift range covered by the simulations including HMXBs (S4 and S5) run with the old version of the code remains negligible. After z∼12 the number begins to rise, demonstrating that the original corrections method could not be used to find the correct values for $\delta T_{\rm b}$ in the QSO only simulations where heating saturation occurs much later.

In conclusion, the original method not only requires two simulations (with and without X-ray sources) but also breaks down when the temperature of the simulation rises above

10^4K and hydrogen atomic cooling becomes efficient. Our new multiphase code is both more computationally efficient and able to correctly calculate X-ray heating in the presence of collisional cooling.

C.3 Final Note

The issues addressed in these Appendices are not unique to C^2-Ray. In fact any simulation not fully resolving the ionized regions will give an incorrect value of δT_b unless temperature saturation has been reached. Eide et al. (2018) have already mentioned our discussion of this problem (see section 3.2.1 in their paper) and are planning to go into more detail in a companion paper. The details of how we have extracted meaningful information from low-resolution simulations and our new algorithm to self consistently track only the temperature of the neutral IGM will certainly be of interest to others facing this problem.